

**ASSESSMENT OF NOAA WATER LEVEL, SEA-SURFACE
TEMPERATURE, AND SALINITY GUIDANCE FROM THE
GLOBAL REAL-TIME OCEAN FORECAST SYSTEM
(G-RTOFS) AND WATER LEVEL GUIDANCE FROM THE
EXTRATROPICAL STORM SURGE (ETSS) SYSTEM IN
WESTERN U.S. COASTAL WATERS**

Silver Spring, Maryland
July 2015



noaa National Oceanic and Atmospheric Administration

U.S. DEPARTMENT OF COMMERCE
National Ocean Service
Coast Survey Development Laboratory

**Office of Coast Survey
National Ocean Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce**

The Office of Coast Survey (OCS) is the Nation's only official chartmaker. As the oldest United States scientific organization, dating from 1807, this office has a long history. Today it promotes safe navigation by managing the National Oceanic and Atmospheric Administration's (NOAA) nautical chart and oceanographic data collection and information programs.

There are four components of OCS:

The Coast Survey Development Laboratory develops new and efficient techniques to accomplish Coast Survey missions and to produce new and improved products and services for the maritime community and other coastal users.

The Marine Chart Division acquires marine navigational data to construct and maintain nautical charts, Coast Pilots, and related marine products for the United States.

The Hydrographic Surveys Division directs programs for ship and shore-based hydrographic survey units and conducts general hydrographic survey operations.

The Navigational Services Division is the focal point for Coast Survey customer service activities, concentrating predominately on charting issues, fast-response hydrographic surveys, and Coast Pilot updates.

**ASSESSMENT OF NOAA WATER LEVEL, SEA-SURFACE
TEMPERATURE, AND SALINITY GUIDANCE FROM THE
GLOBAL REAL-TIME OCEAN FORECAST SYSTEM
(G-RTOFS) AND WATER LEVEL GUIDANCE FROM THE
EXTRATROPICAL STORM SURGE (ETSS) SYSTEM IN
WESTERN U.S. COASTAL WATERS**

Zizang Yang, Philip Richardson, and Richard A. Schmaltz, Jr.
Office of Coast Survey, Coast Survey Development Laboratory,
Silver Spring, MD

July 2015



noaa National Oceanic and Atmospheric Administration

**U. S. DEPARTMENT
OF COMMERCE**
Penny Pritzker,
Secretary

**National Oceanic and
Atmospheric Administration**
Dr. Kathryn D. Sullivan,
Under Secretary

National Ocean Service
Holly Bamford,
Assistant Administrator

Office of Coast Survey
Rear Admiral Gerd F. Glang

Coast Survey Development Laboratory
Captain Richard T. Brennan

NOTICE

Mention of a commercial company or product does not constitute an endorsement by NOAA. Use for publicity or advertising purposes of information from this publication concerning proprietary products or the tests of such products is not authorized.

TABLE OF CONTENTS

LIST OF FIGURES	v
LIST OF TABLES	ix
EXECUTIVE SUMMARY	xi
1. INTRODUCTION	1
1.1. G-RTOFS and ETSS	1
1.2. Observed Data	5
2. METHODS OF MODEL SKILL ASSESSMENT	9
2.1. Model and Data Time Series	9
2.2. Forecast Cycle Based Method	13
2.3. Forecast Hour Based Method	13
3. PERFORMANCE OF WATER LEVEL FORECAST	17
3.1. ETSS	17
3.2. G-RTOFS	26
3.3. Summary	35
4. PERFORMANCE OF G-RTOFS FOR SEA-SURFACE TEMPERATURE	37
4.1. Compared with CO-OPS Observations	37
4.2. Compared with NDBC Buoy Observations	46
4.3. Compared with WOA09 Data	55
4.4. Summary	58
5. PERFORMANCE OF G-RTOFS FOR SEA-SURFACE SALINITY	61
5.1. Forecast Cycle Based Assessment	61
5.2. Forecast Hour Based Assessment	64
5.3. Summary	64
6. SUMMARY	67
ACKNOWLEDGMENTS	69
REFERENCES	69
APPENDIX A. CO-OPS WATER LEVEL STATIONS USED FOR THE ETSS SKILL ASSESSMENT	70
APPENDIX B. CO-OPS WATER LEVEL STATIONS USED IN THE G-RTOFS SKILL ASSESSMENT	71
APPENDIX C. CO-OPS METEOROLOGICAL OBSERVATION STATIONS USED IN THE G-RTOFS SST SKILL ASSESSMENT	72

APPENDIX D. NDBC BUOY STATIONS USED IN THE G-RTOFS SST SKILL ASSESSMENT	73
APPENDIX E. WOA09 GRID NODES FOR ASSESSING THE G-RTOFS SST/SSS	75

LIST OF FIGURES

Figure 1.1. An illustration of the G-RTOFS grid with each cell representing 54 rows and 75 columns of the grid. This grid plot is obtained from http://polar.ncep.noaa.gov/global/about/	2
Figure 1.2. The G-RTOFS forecasts of (a) subtidal water level, (b) SST, and (c) SSS fields at 0300 UTC on 1 October 2012 in the western U.S. coastal waters.....	3
Figure 1.3. The ETSS model guidance areas.....	4
Figure 1.4. Station locations of the ETSS point forecast guidance in the U.S. Pacific Northwest area. This figure is a screen shot from http://www.nws.noaa.gov/mdl/etsurge/index.php?page=map&region=wc&datum=msl&list=&map=0-48&type=&stn=	5
Figure 1.5. Maps of CO-OPS water level stations used for evaluating performances of (a) G-RTOFS and (b) ETSS subtidal water level forecasts.....	6
Figure 1.6. The map of the CO-OPS meteorological observation stations selected for evaluating the performance of G-RTOFS SST forecast.....	7
Figure 1.7. The map of NDBC buoy stations selected for evaluating the performance of G-RTOFS SST forecast.	7
Figure 1.8. A map of 30 grid nodes in WOA09 database. The WOA09 SST and SSS defined these nodes which were used to evaluate the G-RTOFS SST/SSS performance	8
Figure 2.1. Subtidal water level time series of ETSS (blue) and CO-OPS at the CO-OPS station 9443090 (red) during the ETSS forecast cycle on 10/01/2013.....	10
Figure 2.2. Subtidal water level time series of G-RTOFS (blue) and CO-OPS weather station 9410660 (red) during the G-RTOFS forecast cycle on 4/11/2013....	10
Figure 2.3. Sea-surface temperature time series of G-RTOFS (blue) and CO-OPS (red) at the CO-OPS station 9415020 during the G-RTOFS forecast cycle 00z UTC 10/31/2013.....	11
Figure 2.4. Sea-surface temperature time series of G-RTOFS (blue) and NDBC buoy 46247 (red) during the G-RTOFS forecast cycle 00z UTC 10/31/2013.....	11
Figure 2.5. Sea-surface temperature time series of G-RTOFS (blue) and WOA09 Station 13 (red) (Table E.1) during the G-RTOFS forecast cycle on 10/11/2012.	12
Figure 2.6. Sea-surface salinity time series of G-RTOFS (blue) and WOA09 Station 13 (red) (Table E.1) during the G-RTOFS forecast cycle on 10/11/2012.	12
Figure 2.7. Data structure of the two-dimensional model-data time series array	14
Figure 2.8. Forecast hour based daily sub-tidal water level time series of G-RTOFS (blue) and CO-OPS observations (green line and red symbols) from 9/28/2012 to 11/12/2012..	14
Figure 2.9. Forecast hour based daily SST time series of G-RTOFS (blue) and CO-OPS observations (red) from 9/28/2012 to 11/12/2012.	15
Figure 2.10. Forecast hour based daily SSH time series of ETSS (blue) and CO-OPS observations (green line and red symbols) from 9/28/2012 to 11/12/2012	15

Figure 2.11. Forecast hour based daily SST time series of G-RTOFS (blue) and WOA09 (red) during the G-RTOFS forecast cycle on 10/11/2012.	16
Figure 2.12. Forecast hour based daily SSS time series of G-RTOFS (blue) and WOA09 (red) during the G-RTOFS forecast cycle on 10/11/2012.	16
Figure 3.1. Color coded RMSE maps of the ETSS water level forecast in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at twelve CO-OPS water level stations (Table A.1).....	18
Figure 3.2. Means (red squares) and standard deviations (blue error bars) of the ETSS water level forecast errors in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at twelve CO-OPS water level stations (Table A.1).	19
Figure 3.3. The RMSE of the ETSS water level forecast at each forecast hour in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013.	21
Figure 3.4. Color coded RMSE maps of the ETSS water level forecast in October 2012. The four plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, and (d) 96, respectively	22
Figure 3.5. Color coded RMSE maps of the ETSS water level forecast in January 2013. The four plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, and (d) 96	23
Figure 3.6. Color coded RMSE maps of the ETSS water level forecast in April 2013. The four plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, and (d) 96.	24
Figure 3.7. Color coded RMSE maps of the ETSS water level forecast in July 2013. The four plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, and (d) 96.	25
Figure 3.8. Color coded RMSE maps of G-RTOFS water level forecasts in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 18 CO-OPS water level stations (Table B.1).	27
Figure 3.9. Mean (red squares) and standard deviation (blue error bars) of the G-RTOFS water level errors for 2013 at 18 CO-OPS water level stations from: (a) October 2012, (b) January 2013, (c) April 2013, and (d) July.	28
Figure 3.10. Station-averaged root mean squared errors of G-RTOFS water level forecasts in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013, respectively	30
Figure 3.11. Color coded RMSE maps of G-RTOFS water level forecasts in October 2012. The six plots correspond to forecast guidance hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120 and (f) 144, respectively	31
Figure 3.12. Color coded RMSE maps of G-RTOFS water level forecasts in January 2013. The six plots correspond to forecast guidance hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144, respectively	32
Figure 3.13. Color coded RMSE maps of G-RTOFS water level forecasts in April 2013. The six plots correspond to forecast guidance hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144, respectively	33

Figure 3.14. Color coded RMSE maps of G-RTOFS water level forecasts in July 2013. The six plots correspond to forecast guidance hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (d) 144, respectively	34
Figure 4.1. Color coded RMSE maps of the G-RTOFS SST forecast in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 13 CO-OPS weather stations (Table C.1)	38
Figure 4.2. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SST errors in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 13 CO-OPS meteorological observation stations (Table C.1)	39
Figure 4.3. Root mean squared errors (RMSE) of the G-RTOFS SST forecast at forecast hours 1-144 in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013	41
Figure 4.4. Color coded RMSE maps of the G-RTOFS SST forecast in October 2012. The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144	42
Figure 4.5. Color coded RMSE maps of the G-RTOFS SST forecast in January 2013. The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144	43
Figure 4.6. Color coded RMSE maps of the G-RTOFS SST forecast in April 2013. The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144	44
Figure 4.7. Color coded RMSE maps of the G-RTOFS SST forecast in July 2013. The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144	45
Figure 4.8. Color coded RMSE maps of the G-RTOFS SST forecast in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 34 NDBC buoys (Table D.1)	47
Figure 4.9. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SST errors in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 34 NDBC stations (Table D.1)	48
Figure 4.10. Root mean squared errors (RMSE) of the G-RTOFS SST forecast at forecast hours 1-144 in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013	50
Figure 4.11. Color coded RMSE maps of the G-RTOFS SST forecast in October 2012. The model-data comparison was made against SST measured at 34 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144	51
Figure 4.12. Color coded RMSE maps of the G-RTOFS SST forecast in January 2013. The model-data comparison was made against SST measured at 34 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 24, (b) 48,	

(c) 72, (d) 96, (e) 120, and (f) 144	52
Figure 4.13. Color coded RMSE maps of the G-RTOFS SST forecast in April 2013. The model-data comparison was made against SST measured at 34 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144	53
Figure 4.14. Color coded RMSE maps of the G-RTOFS SST forecast in July 2013. The model-data comparison was made against SST measured at 34 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144	54
Figure 4.15. Color coded RMSE maps of the G-RTOFS SST forecast in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 28 WOA09 data grid points (Table E.1).	56
Figure 4.16. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SST errors in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 28 WOA09 data grid points (Table E.1)	57
Figure 4.17. Root mean squared errors (RMSE) of the G-RTOFS SST forecast at forecast hours 1-144 in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 28 WOA09 data grid points (Table E.1).....	59
Figure 5.1. Color coded RMSE maps of the G-RTOFS SSS forecast in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 28 WOA09 data grid points (Table E.1)	62
Figure 5.2. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SSS errors in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 28 WOA09 data grid points (Table E.1)	63
Figure 5.3. Root mean squared error (RMSE) of the G-RTOFS SSS forecast at forecast hours 1-144 in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 28 WOA09 data grid points (Table E.1).....	65

LIST OF TABLES

Table 3.1. Forecast cycle-based statistics of the G-RTOFS and ETSS SSH errors.....	36
Table 3.2. Forecast hour-based RMSE of the G-RTOFS and ETSS SSH	36
Table 4.1. Forecast cycle-based statistics of the G-RTOFS SST errors	60
Table 4.2. Forecast hour-based RMSE of the G-RTOFS SST..	60

EXECUTIVE SUMMARY

The Operational Nowcast/Forecast Systems (OFS), presently being developed by the National Ocean Service (NOS) of the National Oceanic and Atmospheric Administration (NOAA), make use of sea-surface height (SSH), sea-surface temperature (SST), and sea-surface salinity (SSS), forecast guidance from the Global Real-Time Ocean Forecast System (G-RTOFS), and SSH forecast guidance from the Extra-Tropical Storm Surge (ETSS) system to provide the open ocean boundary forcings. To support future development of NOS OFS in western U.S. coastal waters, we assessed the performance of the G-RTOFS forecasts for SSH, SST, and SSS as well as the performance of the ETSS forecasts for SSH.

The G-RTOFS operates on four six-hourly cycles (00z, 06z, 12z, and 18z UTC) each day, while ETSS operates on a daily cycle (00z UTC). This study focusses on analyzing the output produced on cycle 00z UTC. The three-hourly G-RTOFS SSH, SST, and SSS forecast was analyzed from hour 00 to hour 144 UTC. The hourly ETSS SSH forecast was analyzed from hour 00 to hour 96 UTC.

Our intention is to gain insight into the model performance from two perspectives: (1) the model performance across a forecast cycle (FC) and (2) the evolution of performance associated with the forecast hour (FH) of a forecast cycle. Accordingly, we developed a FC-based method and a FH based method, which are described in Chapter 2. We applied the FC method to estimate the bias, standard deviation, and root-mean-squared error of a forecast cycle over a series of cycles. The FH method was applied to estimate the root-mean-squared error for each given forecast hour of the forecast cycle over a series of cycles.

We conducted model-data comparisons over four separate month long periods: October 2012, and January, April, and July of 2013. We used these periods as proxies for fall, winter, spring, and summer, respectively. In this way, the analysis covers all four seasons of the year.

With regard to observed data, we used quality-controlled hourly water level measurements from 18 CO-OPS water level stations (Table A.1 and Table B.1). We extracted the subtidal water level values by low-pass filtering the total water level time series which were then used to assess the performance of the model SSH.

We used SST measurements from 13 CO-OPS meteorological observation stations and 34 NDBC buoys to evaluate the performance of the G-RTOFS SST forecast. In general the NDBC stations are located further offshore than are the CO-OPS stations. Hence, the NDBC data are more representative of the offshore SST conditions. In addition, we used SST/SSS from 30 coastal grid points (Table E.1) in the National Oceanographic Data Center's World Ocean Atlas 2009 (WOA09) database to assess the agreement of the G-RTOFS forecast to climatology.

In general, the forecast guidance of SSH (from G-RTOFS and ETSS) and of SST/SSS (from G-RTOFS) demonstrates satisfactory agreement with observations. The performance skill varies by season and evolves as the forecast hour progresses through the forecast cycle.

For SSH from both models, the standard deviation (STD) with respect to the forecast cycle ranges from around 2 to 4 cm throughout all four seasons. The bias of ETSS is about -7 cm in fall and winter and about 0.1 to 2.2 cm in spring and summer. In contrast, G-RTOFS displays a bias of 3 to 4 cm over all four seasons. The root mean squared error of ETSS ($RMSE_{ETSS}$) increases from 6 to 7 cm in spring and summer to ~9 cm in fall and winter. The RMSE for G-RTOFS increases from ~7 cm in summer and fall to around 11 to 14 cm in winter and spring.

With respect to the forecast hour, the skill of both models degrades as the forecast hour progresses. The $RMSE_{ETSS}$ increases from less than 2 cm at the beginning of a forecast cycle to 6-10 cm near the end of the cycle. The root-mean-squared error of G-RTOFS first increases monotonically from less than 2 cm at the start of a forecast cycle, then asymptotically approaches a plateau value. This plateau value is maintained roughly from mid-cycle through the end of cycle. Model skill also demonstrates seasonal variability. The specific plateau value and the hour of transition vary by month. The maximum $RMSE_{ETSS}$ is 8 cm, 10 cm, 10 cm, and 6 cm in fall, winter, spring, and summer, respectively. The maximum $RMSE_{G-RTOFS}$ is 6.5 cm, 11.5 cm, 10.0 cm, and 4.7 cm in fall, winter, spring, and summer, respectively. The RMSE for both models is lowest in the summer.

The performance of the G-RTOFS SST forecast demonstrates appropriate seasonal variability. The station-averaged bias, STD, and RMSE remain on similar orders of magnitudes in the fall, winter, and summer, and are noticeably greater in the summer. For instance, the bias ranges from around 1.9 to 3.7 °C in the summer and is less than 1 °C in the other seasons. RMSE is about 2.2 to 3.8 °C in summer and less than 1.3 °C in the other seasons. The STD exhibits smaller seasonal variability, though its summer value is still about 0.1 to 1.1 °C greater than the average magnitude of the other seasons. The model skill does not degrade with the progressing forecast hour of the forecast cycle. The RMSE remains nearly invariable throughout each cycle as opposed to the gradual skill degradation in the case of the G-RTOFS SSH forecast.

The performance of the G-RTOFS SSS forecast demonstrates a satisfactory agreement with the WOA09 database. The station-averaged bias ranges from between -0.3 psu and 0.3 psu in the fall, winter, and spring and spreads over a slightly broader range from between -0.5 psu and -0.3 psu in summer. The STD is less than 0.2 psu throughout all four seasons. The RMSE appears to be nearly constant (~0.2 psu) throughout the forecast cycle across all four seasons.

1. INTRODUCTION

The National Ocean Service (NOS) of NOAA has been developing operational forecast systems (OFS) to produce nowcast/forecast guidance of ocean state variables including water levels, temperature (T), salinity (S), and three-dimensional (3-D) currents in the U.S. estuaries, coastal, and shelf waters. The OFSs produce valuable information to support safe maritime navigation, emergency response, and coastal environment management. The backbone of the various systems are the hydrodynamic modeling systems which are forced with water levels, temperature, salinity, and currents on the model domain's open ocean boundary, as well as the meteorological forcing on the surface and the river discharge at the river entrance. Open boundary forcing plays a critical role in the accuracy of the OFS nowcast/forecast guidance.

The NOAA National Weather Service (NWS) Global Real-Time Ocean Forecast System (G-RTOFS) and Extra-Tropical Storm Surge (ETSS) Model are two operational systems operated by the NOAA's National Center for Environmental Predictions (NCEP). The NOS OFSs normally use subtidal water level and 3-dimensional (3-D) T/S data from G-RTOFS to drive their hydrodynamic model runs and use the ETSS subtidal water level output as backup when the G-RTOFS water level output is not available. Hence, it is worthwhile to evaluate the performance of the two models. As a first step in model evaluation, the present project focuses on assessing the G-RTOFS water level skills and SST/SSS skills, and the ETSS water level skills.

In the present project, we aim to assess the ETSS and G-RTOFS skills in the U.S. western coastal waters and we focus on four months: October 2012, and January, April, and July of 2013. These months roughly correspond to the seasons of fall, winter, spring, and summer, respectively. We evaluated the model performance by comparing the model results with in-situ observations (for water level and SST) as well as with data from the climatological monthly world ocean database (for SST and SSS).

The remainder of Section 1 introduces the background information about setups and operations of both G-RTOFS and ETSS as well as observational data sets used in this study. Section 2 describes technical details on two methods, the forecast cycle (FC) based method and the forecast hour (FH) based method, used to assess the model performance. Sections 3-5 discuss the model performance of SSH (for G-RTOFS and ETSS), SST (for G-RTOFS), and SSS (for G-RTOFS) guidance, respectively. Section 6 summarizes the model assessment results and the methods used to attain those results.

1.1. G-RTOFS and ETSS

This section gives an introduction to the model setup and operations of both G-RTOFS and ETSS.

1.1.1. G-RTOFS

G-RTOFS is based on the Naval Oceanographic Office's (NAVO) configuration of the $1/12^\circ$ eddy resolving global Hybrid Coordinates Ocean Model (HYCOM) (G-RTOFS, 2011). It is initialized daily with NAVO generated initial conditions using the Navy Coupled Ocean Data

Assimilation (NCODA) system (Metzger, etc., 2005). The system assimilates in situ profiles of temperature and salinity from a variety of sources and remotely sensed SST, SSH and sea-ice concentrations. G-RTOFS is forced with 3-hourly momentum, radiation, and precipitation fluxes from the operational Global Forecast System (GFS) fields.

The G-RTOFS ocean model has 32 vertical hybrid layers (isopycnals in the deep, isolevel (Z-level) in the mixed layer, and sigma in shallow waters) and a horizontal grid of dimension 4500×3298 . The grid has an Arctic bi-polar patch north of 47°N and a Mercator projection south of 47°N through 78.6°S (Figure 1). The coastline is fixed at 10-m isobaths with the Bering Straits being open. The potential temperature is referenced to 2000 m depth (sigma-2) and the first level is fixed at 1 m depth.

G-RTOFS became operational at the NWS NCEP Environmental Modeling Center (EMC) on 25 October 2011. It runs once a day and produces, at three hour intervals, forecast output for sea surface values (SSH, SST, and SSS) and at six hour intervals, forecast output for full volume parameters (3-dimensional temperature, salinity, currents, and mixed layer depths) from the initial time of (0Z) and extending out to 144 hours. The operation period extended from 25 October 2011 through 25 July 2013. On 25 July 2013, the length of the forecast period was extended from 144 hours to 192 hours. G-RTOFS has run at NCEP, with this extended forecast period, ever since.

As sample illustrations, Figures 1.2(a-c) display a snapshot of the G-RTOFS forecasts of (a) subtidal water level, (b) SST, and (c) SSS fields in an area along the U.S. west coast at 0300 UTC on 1 October 2012.

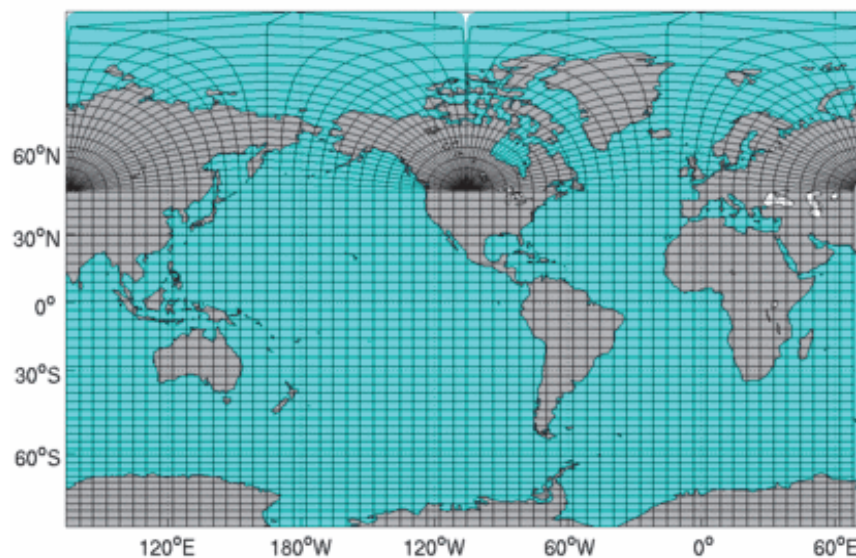


Figure 1.1. An illustration of the G-RTOFS grid with each cell representing 54 rows and 75 columns of the grid. This grid plot is obtained from <http://polar.ncep.noaa.gov/global/about/>.

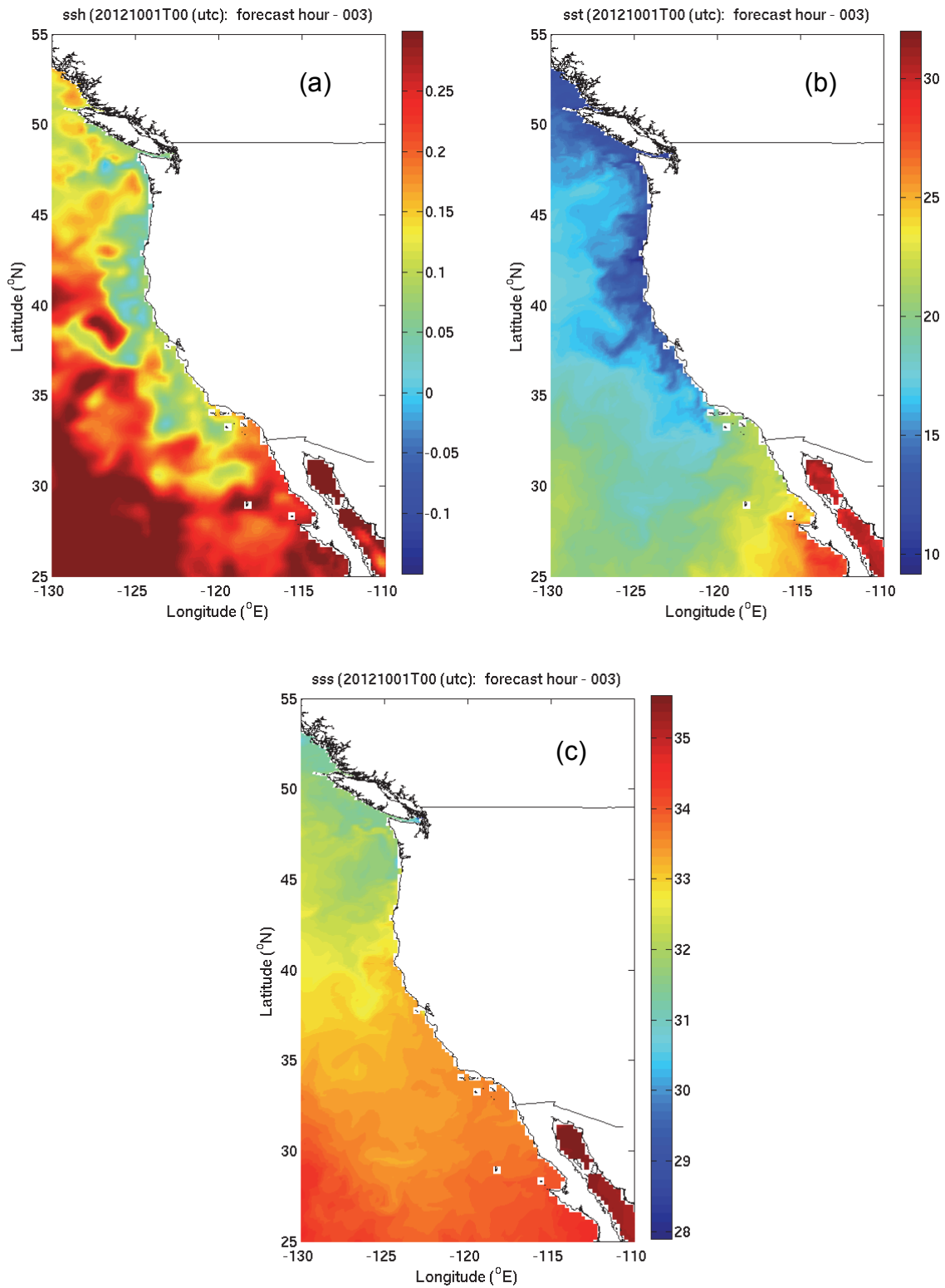


Figure 1.2. The G-RTOFS forecasts of (a) subtidal water level, (b) SST, and (c) SSS fields at 0300 UTC on 1 October 2012 in the western U.S. coastal waters.

1.1.2. ETSS

The ETSS model was developed by the Meteorological Development Laboratory (MDL) of the NWS. It is a variation on the NWS's Sea, Lake and Overland Surges from Hurricanes (SLOSH) model (ETSS, 1992). It is a prognostic, two-dimensional, barotropic model forced by real time output of winds and air pressures from the NCEP GFS. ETSS has been applied in multiple areas in the U.S. Arctic, Pacific, Atlantic, and the Gulf of Mexico continental shelf and coastal waters. Figure 1.3 displays a map of the ETSS guidance areas in the Arctic and Gulf of Alaska, the U.S. Northwest, the Gulf of Mexico, and the Southeast, Mid-Atlantic, and Northeast coasts of the U.S.

ETSS initially ran operationally on the NCEP's Central Computing System (CCS) four times daily out to 96 hours producing numerical storm surge guidance (subtidal water levels) for extra-tropical systems. Beginning 26 August 2012, the ETSS operation was transitioned onto the NCEP's Weather and Climate Operational Supercomputer Systems (WCOSS).

NCEP disseminates the ETSS forecast guidance of subtidal water levels in both point and gridded formats. We used the point guidance outputs in the present study. Figure 1.4 illustrates the station locations in the U.S. Pacific Northwest coastal waters. Since this study focuses on evaluating the ETSS performance in open coastal waters, some estuarine stations are discarded. Table A.1 lists the names and geographical locations of the stations selected for model-data comparisons.

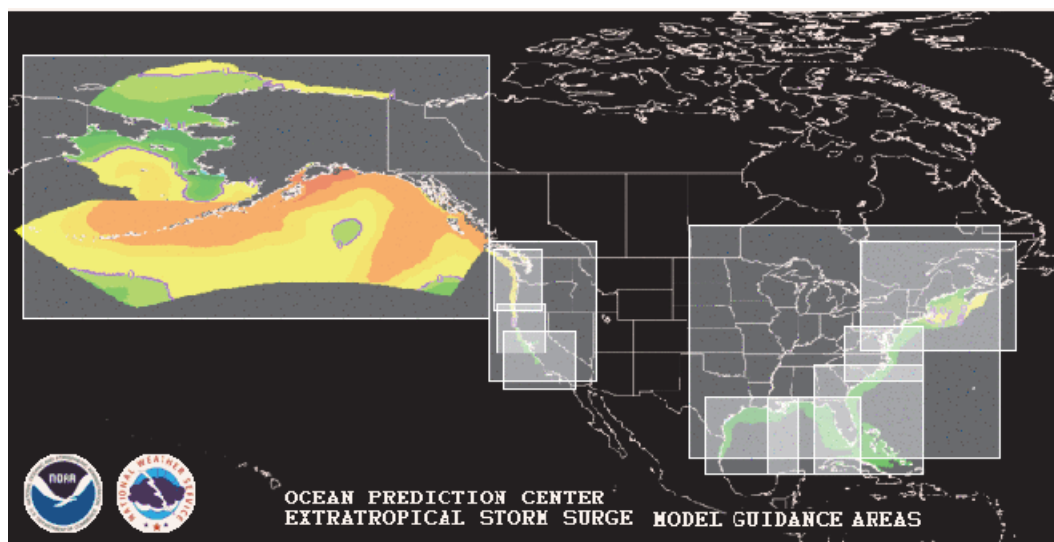


Figure 1.3. The ETSS model guidance areas.

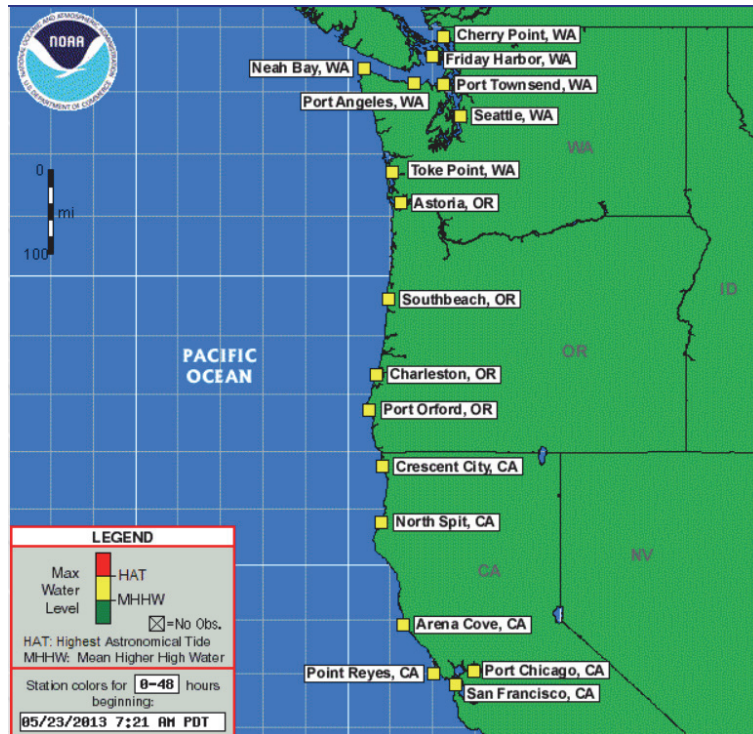


Figure 1.4. Station locations of the ETSS point forecast guidance in the U.S. Pacific Northwest area. This figure is a screen shot from <http://www.nws.noaa.gov/mdl/etsurge/index.php?page=map®ion=wc&datum=msl&list=&map=0-48&type=&stn=>.

1.2. Observed Data

We used observational data from three sources: (1) water level and sea-surface temperature (SST) collected at water level and meteorological observation stations of NOAA's Center for Operational Oceanographic Products and Services (CO-OPS), (2) SST from the NOAA's National Data Buoy Center's (NDBC) buoy measurements, and (3) SST and sea-surface salinity (SSS) from the 2009 version of the National Oceanographic Data Center's (NODC's) World Ocean Atlas (WOA09) database. The following sections describe in detail about the data sources and retrievals.

1.2.1. CO-OPS Water Level

Figures 1.5(a) and (b) display maps of CO-OPS water level stations at which we evaluated the performance of G-RTOFS and ETSS forecasts, respectively. Since G-RTOFS and ETSS both focus on forecasts in the open ocean and coastal areas, we purposefully selected only offshore stations for the model-data comparison and excluded those located in estuaries or embayments.

We used quality-controlled hourly total water level time series from CO-OPS for model-data comparisons. The total water level is a supposition of both tidal and non-tidal components. To evaluate the G-RTOFS and ETSS non-tidal surface elevations, we first applied a low-pass filter

to the total water level time series to extract the subtidal water levels. The filtered time series were then compared with the G-RTOFS and ETSS results.

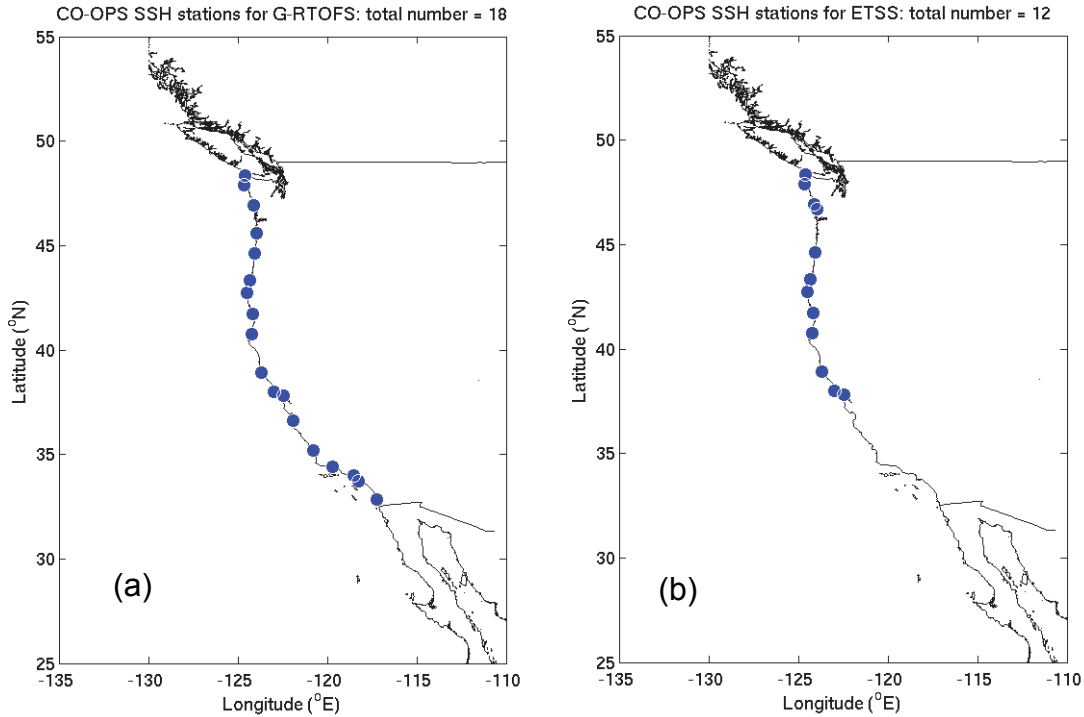


Figure 1.5. Maps of CO-OPS water level stations used for evaluating performances of (a) G-RTOFS and (b) ETSS subtidal water level forecasts.

1.2.2. CO-OPS SST

We used hourly SST time series collected at the CO-OPS meteorological observation stations. Figure 1.6 illustrates locations of a total 13 stations employed in the present study. Table C.1 lists the station identification (ID) numbers, names, and their geographical locations in longitude and latitude. We downloaded hourly SST time series data from the CO-OPS online database.

1.2.3. NDBC SST

We used SST measurements from 34 NDBC buoy measurements along the U.S. West Coast. Table D.1 lists the buoy IDs, station names, station locations in longitude and latitude. Figure 1.7 displays a map of these stations. We downloaded the hourly SST time series data from <http://www.ndbc.noaa.gov/data/realtime2>.

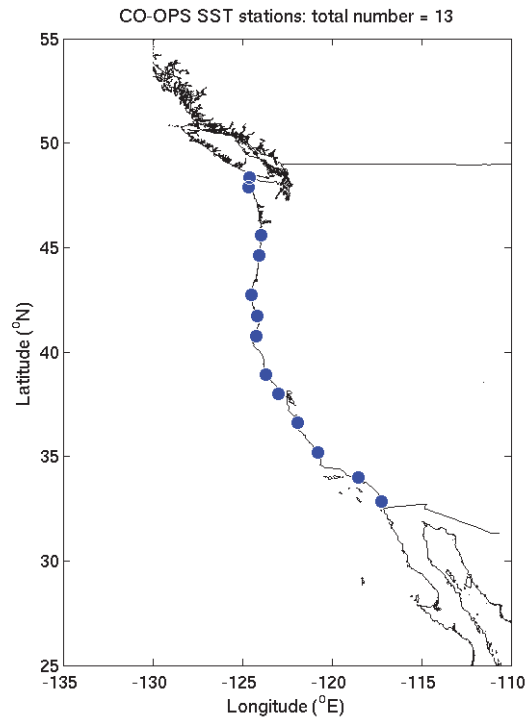


Figure 1.6. The map of the CO-OPS meteorological observation stations selected for evaluating the performance of G-RTOFS SST forecast.

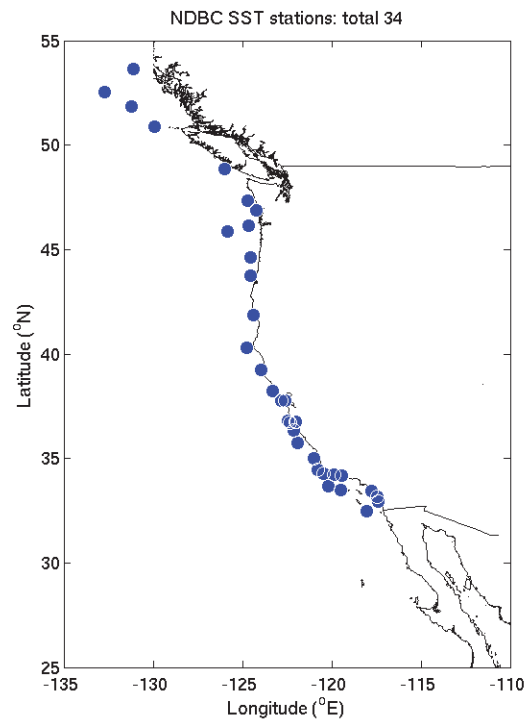


Figure 1.7. The map of NDBC buoy stations selected for evaluating the performance of G-RTOFS SST forecast.

1.2.4. WOA09 SST and SSS

The World Ocean Atlas 2009 (WOA09) is a set of objectively analyzed climatological fields with a grid of 1° resolution. The climatological fields include in situ temperature, salinity, dissolved oxygen, apparent oxygen utilization (AOU), percent oxygen saturation, phosphate, silicate, and nitrate at standard depth levels for annual, seasonal, and monthly compositing periods for the World Ocean. It also includes associated statistical fields of observed oceanographic profile data interpolated to standard depth levels on both 1° and 5° grids. We downloaded the data set in NetCDF format from the NODC Web site: http://www.nodc.noaa.gov/OC5/WOA09F/pr_woa09f.html.

In the WOA09 database, quantities defined on each grid node represent a composite of observations collected in nearby areas. In the present study, we treat each grid node as an in situ observing station. We selected 28 nearshore grid nodes (Figure 1.8) to form a station inventory for model-data comparisons. Table E.1 in Appendix E lists the geographical position by longitude and latitude for each selected grid node.

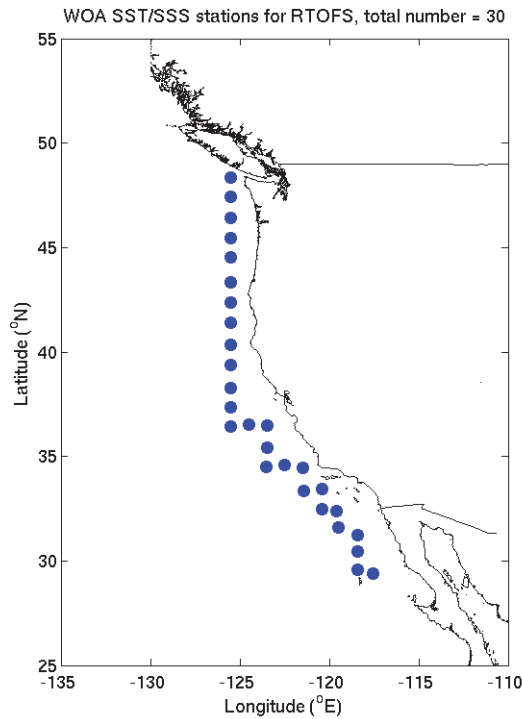


Figure 1.8. A map of 28 grid nodes from the WOA09 database. The WOA09 SST and SSS defined these nodes which were used to evaluate the G-RTOFS SST/SSS performance.

2. METHODS OF MODEL SKILL ASSESSMENT

We evaluated the performance of G-RTOFS forecast guidance on subtidal water level (SSH), sea-surface temperature (SST), and sea-surface salinity (SSS) as well as the performance of the ETSS (SSH) forecast guidance. We assessed the model performance from two perspectives: forecast cycle-based (FC) performance and forecast hour-based (FH) performance. The former evaluates the performance over the forecast cycle as a whole, whereas the latter evaluates the performance at individual hours within the forecast cycle.

G-RTOFS produces one forecast cycle each day, at hour 00 UTC, which produces 3-hourly forecast guidance up to 144 or 192 hours (Section 1.2). In the present study, we focus on the forecast through 144 hours. Hence, every forecast cycle produces a 49 point time series of SSH, SST, and SSS corresponding to hours 00, 03, 06, ..., 144 relative to hour 00 UTC of that day.

The ETSS water level forecast is run on hours 00, 06, 12, and 18, UTC. Each of the four forecast cycles produces an hourly forecast up to 96 hours. Since all four cycles utilize the same model configuration (grid) and forcing data of the same quality, it is reasonable to believe that the model will perform equally well across all four cycles. For this reason, we focus on evaluating the model performance of the 00 UTC cycle. In this context, we refer to the ETSS forecast hours as 01, 02, 03, ..., 96 UTC.

For each variable (SSH, SST, and SSS) at a given station, we first compile two coincident time series. The first is from the model output and the other is from the observed data. Data points in the two time series correspond to the same model forecast guidance hour. For each pair of values, model and observed, we subtract the observed value from the model value to get a time series representing the model-data difference. We then apply the FC and FH methods to this new time series to assess the model performance.

Sections 2.2 and 2.3 describe details of the FC and FH methods, respectively. Since the methods used for water level, SST, and SSS are the same, in the following sections we will differentiate the variable names only if necessary.

2.1. Model and Data Time Series

To compare the forecast with observations, we first identify the model grid nodes collocated with the observed data stations. The in-situ data always possess a higher sampling rate (usually hourly) than model series (three hourly for G-RTOFS and hourly for ETSS). Consequently, we subset the in-situ data to form an observed data time series coincident with the model time series.

We create the model-data difference time series by subtracting each observed data value from the corresponding model value. Since we utilize one forecast cycle, the forecast cycle generated at hour 00 UTC, we ended up with time series for each day. For G-RTOFS, each time series consists of 49 data points corresponding to the three-hourly forecast guidance from hours 00 through 144 UTC. For ETSS, each time series consists of 96 data points corresponding to the hourly forecast guidance from hours 00 through 96. In a given month, this results in 30 (for

April) or 31 (for January, July, and October) time-series which form the base-line data sets for further data analysis.

Figures 2.1 - 2.6 display sample model and observed time series. Specifically, Figure 2.1 depicts the ETSS and CO-OPS subtidal water levels. The step function like appearance of the ETSS forecast is due to the data accuracy being up to 0.1 inch. Figure 2.2 depicts the G-RTOFS and CO-OPS subtidal water levels. Figure 2.3 depicts the G-RTOFS and CO-OPS SST. Figure 2.4 depicts the G-RTOFS and NDBC buoy SST. Finally, Figures 2.5 and 2.6 display the G-RTOFS and WOA09 SST and SSS, respectively.

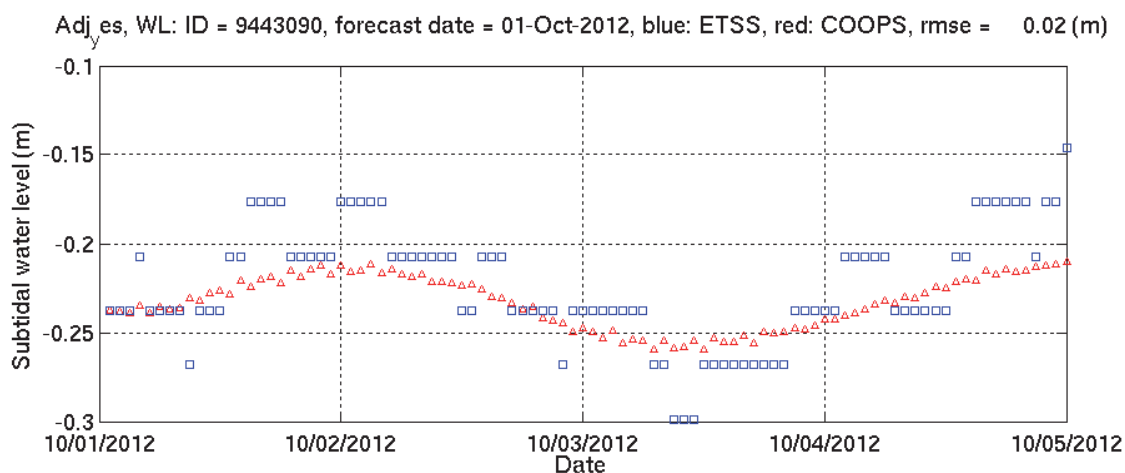


Figure 2.1. Subtidal water level time series of ETSS (blue) and CO-OPS at the CO-OPS station 9443090 (red) during the ETSS forecast cycle on 10/01/2013.

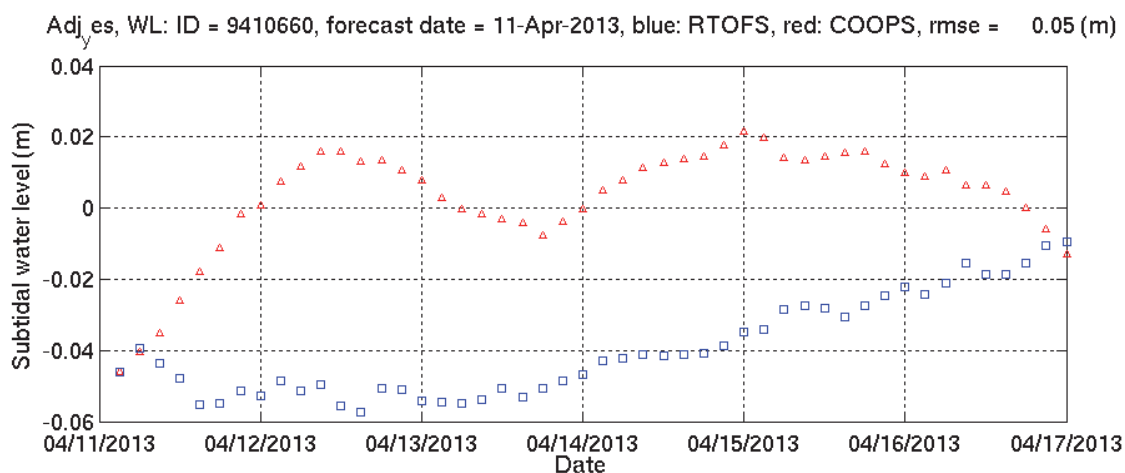


Figure 2.2. Subtidal water level time series of G-RTOFS (blue) and CO-OPS weather station 9410660 (red) during the G-RTOFS forecast cycle on 4/11/2013.

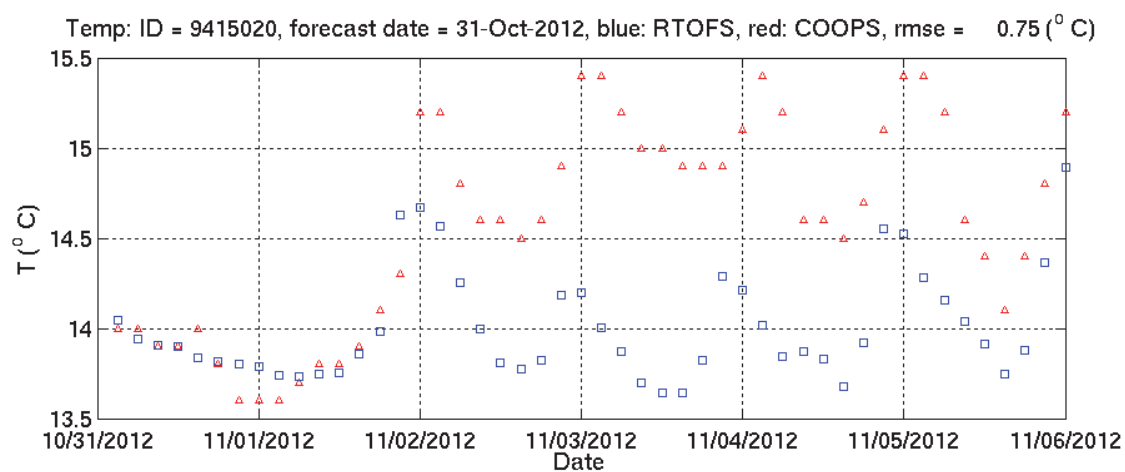


Figure 2.3. Sea-surface temperature time series of G-RTOFS (blue) and CO-OPS (red) at the CO-OPS station 9415020 during the G-RTOFS forecast cycle 00z UTC 10/31/2013.

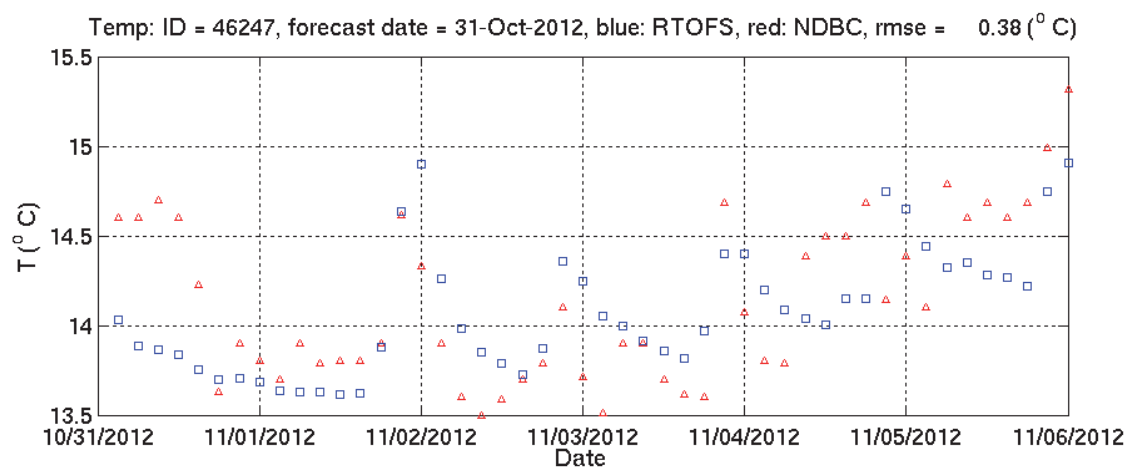


Figure 2.4. Sea-surface temperature time series of G-RTOFS (blue) and NDBC buoy 46247 (red) during the G-RTOFS forecast cycle 00z UTC 10/31/2013.

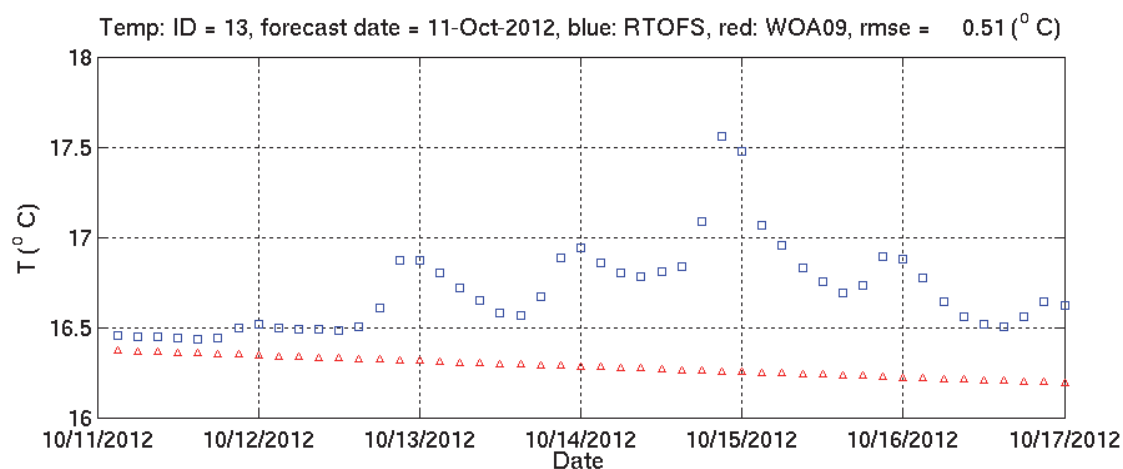


Figure 2.5. Sea-surface temperature time series of G-RTOFS (blue) and WOA09 Station 13 (red) (Table E.1 in Appendix E) during the G-RTOFS forecast cycle on 10/11/2012.

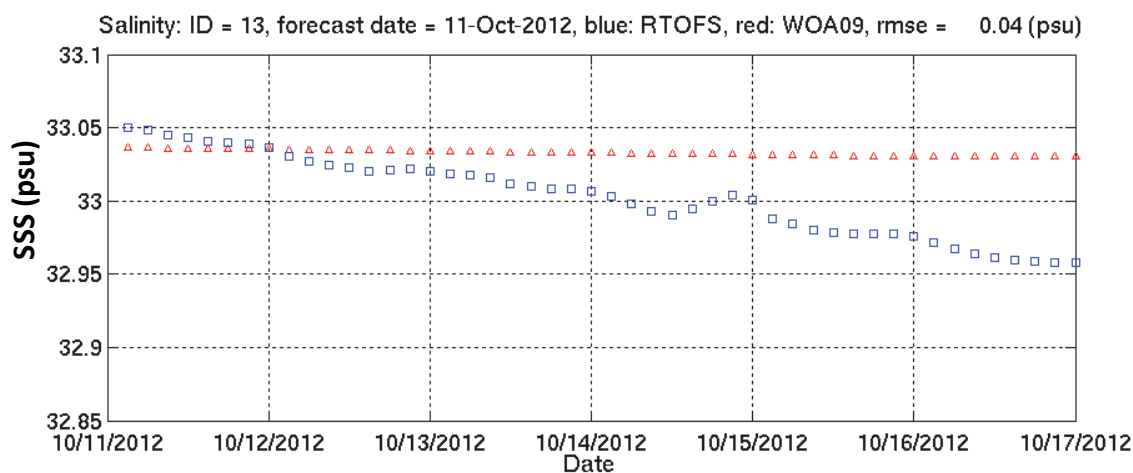


Figure 2.6. Sea-surface salinity time series of G-RTOFS (blue) and WOA09 Station 13 (red) (Table E.1 in Appendix E) during the G-RTOFS forecast cycle on 10/11/2012. Note that the WOA09 database has only the monthly averaged values. Its plot (red triangles) was based on the linear interpolation between the October and November datasets.

2.2. Forecast Cycle Based Method

This method evaluates the model performance across the entire forecast cycle. From each model-data difference time series (Section 2.1), we calculate the mean, standard deviation (STD), and the root-mean-square error (RMSE) of the model forecast compared with the observed data.

One model-data time series is created for each forecast cycle (Section 2.1). Since one forecast cycle is utilized per day, either 30 or 31 model-data difference time series are created for each month. For each of these time series, we calculate the daily mean, STD, and RMSE. Hence, we obtain 30 or 31 values for each statistical parameter for each month. Taking the average of these values, for each parameter, gives us a monthly mean value for mean, STD, and RMSE.

2.3. Forecast Hour Based Method

This method evaluates the model performance at a given forecast hour across multiple forecast cycles. It may reveal how the model skill evolves with the increasing of the forecast guidance hour.

Figure 2.7 illustrates the scheme of this method. As described in Section 2.1, we compile either 30 or 31 model-data pairs for each given month, with each pair corresponding to a time series from one forecast cycle. This essentially forms a two-dimensional array; one dimension is the forecast guidance hour and the other is the daily forecast cycle. The FH method focuses on contrasting model-data discrepancies at each forecast guidance hour. We therefore subset the 2-D array into multiple 1-D time series according to the forecast guidance hour, i.e., each 1-D series consists of data points corresponding to the same forecast hour, however, from multiple forecast cycles. The number of pairs of data points in each time series is either 30 or 31 depending on the number of days (i.e., forecast cycles) in the month under consideration.

ETSS produces, for each forecast cycle, hourly SSH guidance from hours 01 to 96 UTC. We therefore compile 96 time series for SSH. G-RTOFS produces forecast guidance at 3 hour intervals from hour 00 to hour 144 UTC. Thus, we have 49 time series of SSH, SST, and SSS corresponding to each forecast hour.

The FH based method was created to investigate the variation of model skill with the increasing forecast hour. To characterize the trend of variation, we offset the model time series SSH by the amount of the first data point minus the first model point. We then studied the model-data discrepancy in the remaining hours. Note that with or without the offsetting adjustment, there is not a noticeable difference in the SST or SSS assessment. Therefore, we only apply the offset to the SSH analysis.

Figures 2.8- 2.12 display the forecast hour based time series. Specifically, Figure 2.8 depicts the G-RTOFS and CO-OPS subtidal water levels; Figure 2.9, the G-RTOFS and CO-OPS SST; Figure 2.10, the ETSS and CO-OPS SSH; Figure 2.11, the G-RTOFS and WOA09 SST; and Figure 2.12 displays the G-RTOFS and WOA09 SSS.

Following the same approach as was taken with the FC method, we estimate the mean, STD, and RMSE of model-data differences at each station. We then average each quantity over 30 or 31 days to derive station-averaged model skills.

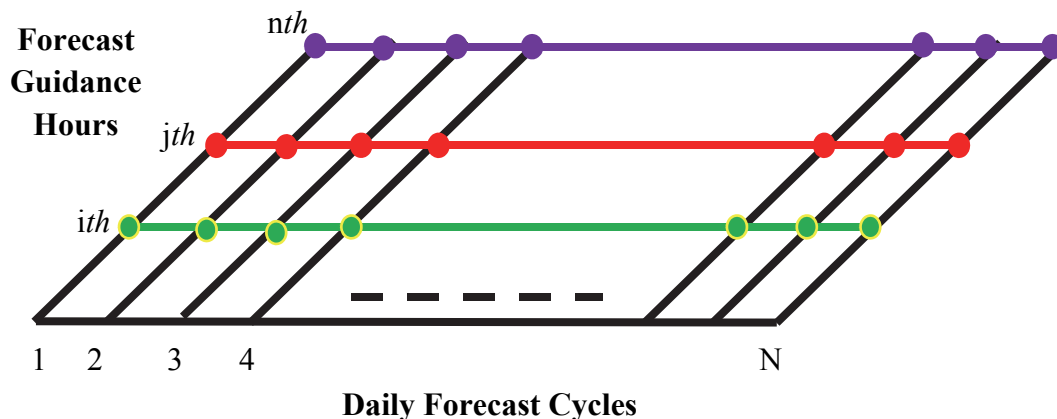


Figure 2.7 Data structure of the two-dimensional model-data time series array.

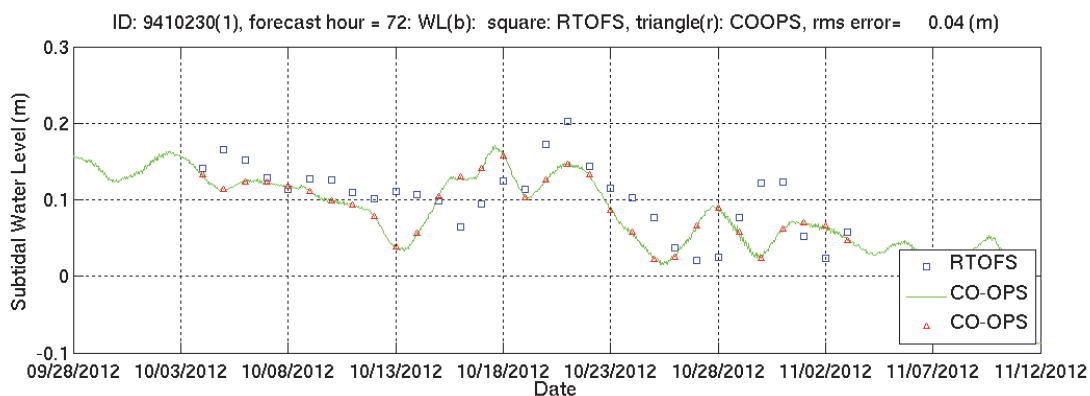


Figure 2.8. Forecast hour based daily sub-tidal water level time series of G-RTOFS (blue) and CO-OPS observations (green line and red symbols) from 9/28/2012 to 11/12/2012.

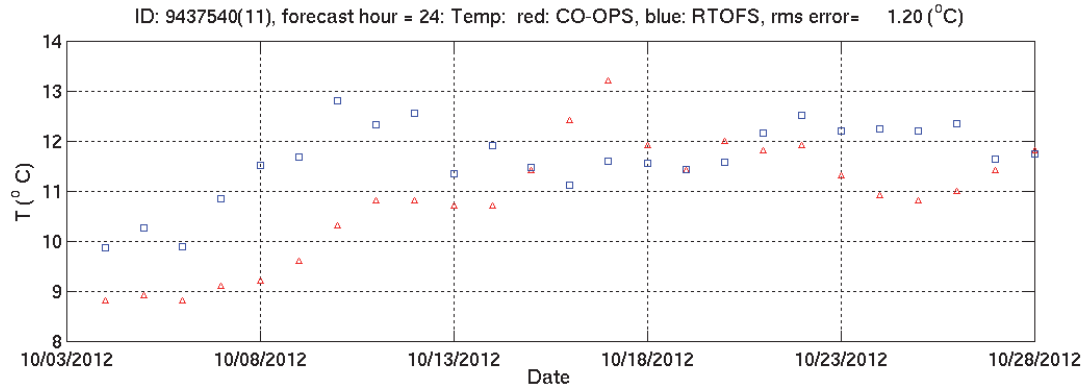


Figure 2.9. Forecast hour based daily SST time series of G-RTOFS (blue) and CO-OPS observations (red) from 10/03/2012 to 10/28/2012.

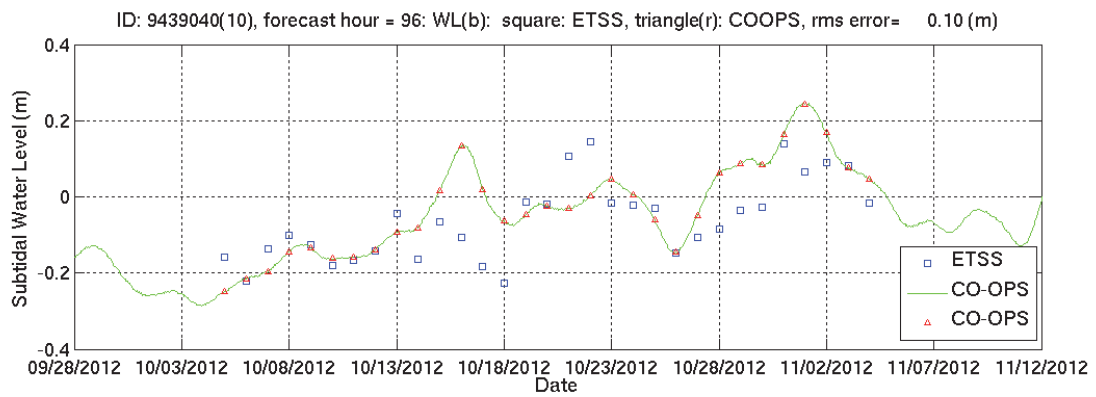


Figure 2.10. Forecast hour based daily SSH time series of ETSS (blue) and CO-OPS observations (green line and red symbols) from 9/28/2012 to 11/12/2012.

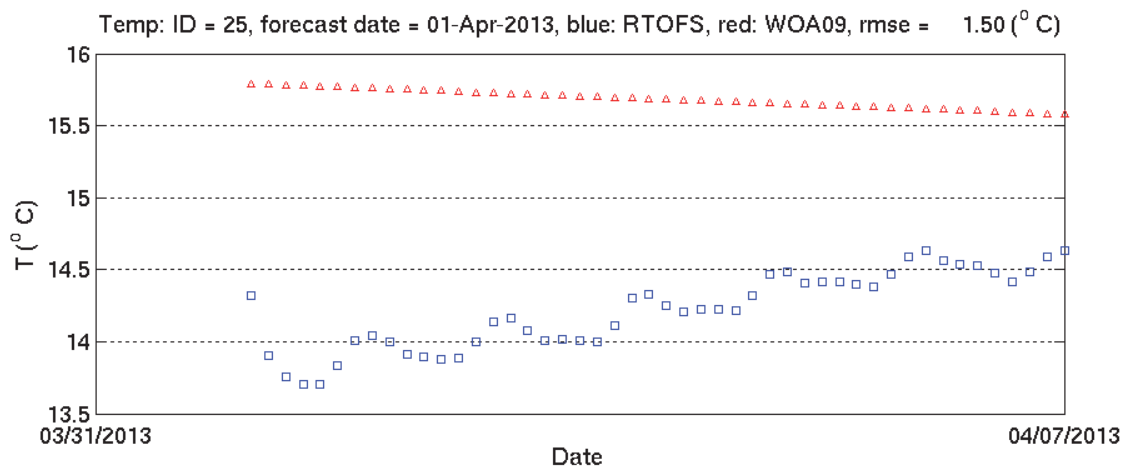


Figure 2.11. Forecast hour based daily SST time series of G-RTOFS (blue) and WOA09 (red) during the G-RTOFS forecast cycle on 10/11/2012.

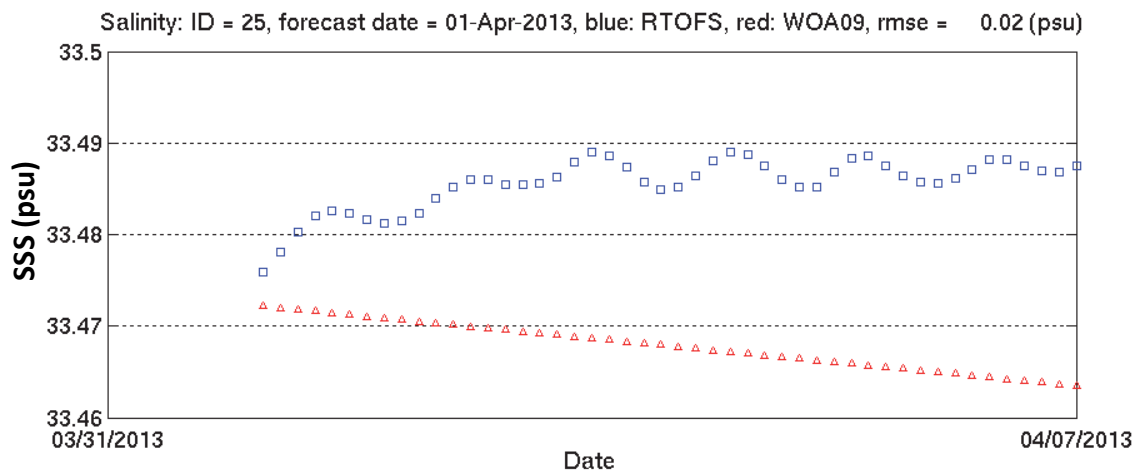


Figure 2.12. Forecast hour based daily SSS time series of G-RTOFS (blue) and WOA09 (red) during the G-RTOFS forecast cycle on 10/11/2012.

3. PERFORMANCE ON WATER LEVEL FORECAST

In this chapter, we evaluate the performance of both ETSS and G-RTOFS for SSH. We apply both the forecast cycle (FC) method and the forecast hour (FH) method in the analysis.

3.1. ETSS Forecast

In the following, we will compare ETSS point guidance outputs with measurements from 12 CO-OPS water level stations (Table A.1). Table A.1 lists the station meta data including their identification (ID) numbers, names, and geographical locations in longitude and latitude.

3.1.1. FC Based Assessment

Figures 3.1(a) - (d) display maps of the model root-mean-square error (RMSE) in October 2012, and in January, April, and July 2013. The RMSE typically ranges between 6 and 13 cm with slight variations in different months. It is around 6 - 11 cm in April and July and displays slightly larger (up to about 7 cm) spatial variations in October and January. This indicates that ETSS had better skill in spring and summer than in fall and winter.

Figures 3.2(a) - (d) display the mean and standard deviation (STD) of model error at the 12 water level stations (Table A.1) in October 2012, January, April, and July 2013. In October and January, the means ranged between 0 and -10 cm, whereas the STD was 2-4 cm. The negative mean indicates that the model underpredicted water levels. In April and July, the mean was 2.2 cm and 0.1 cm, respectively, whereas the corresponding STD was 3.9 cm and 2.6 cm.

3.1.2. FH Based Assessment

In this section, we apply the FH method to evaluate the model performance for SSH at individual hours within a forecast cycle.

Station-Averaged RMSE Figure 3.3 displays the station averaged RMSE of ETSS water levels at each forecast hour in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013. As anticipated, model skill gradually degrades with the increasing forecast hour; the RMSE increased from 1-2 cm at hour 2 to 6-10 cm at hour 96.

The ranking of model performance by month from best to worst is July 2013 and October 2012, followed by a nearly equal April and January of 2013. For instance, at a time in about the middle of a forecast cycle, the RMSE in July, October, April, and January was about 4, 5, 6, and 7 cm, respectively.

Note that the RMSE displays a rather regular sinusoidal oscillation with a period of about 12 hours. This does not reflect any inherent characteristics of the model skill, but comes from an alias due to the low sampling rate, daily in this case, of the underlying time series.

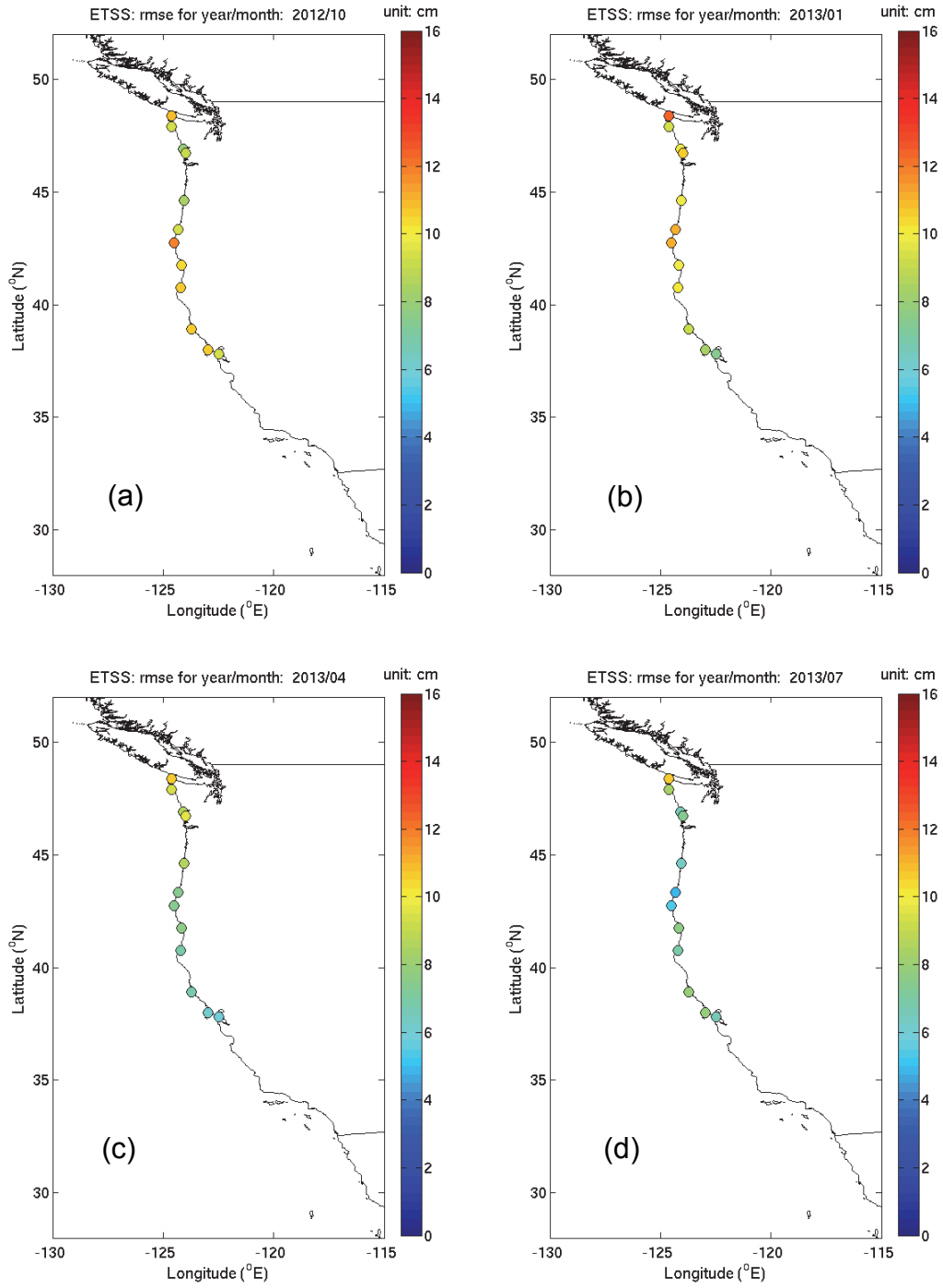
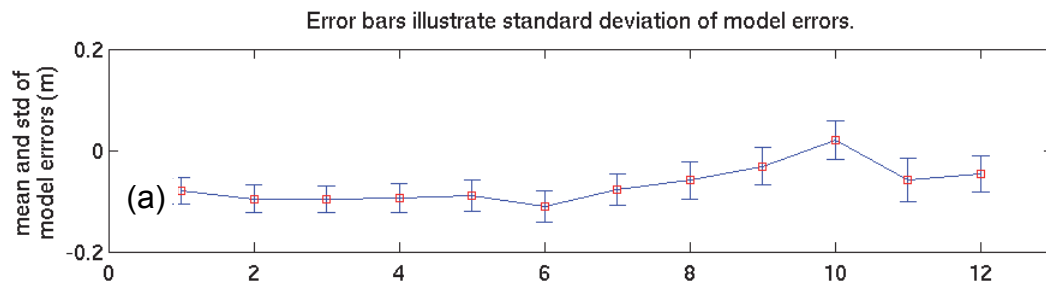
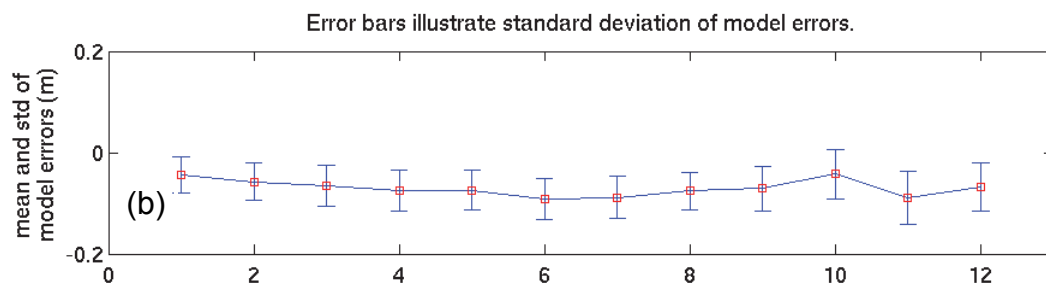


Figure 3.1. Color coded RMSE maps of the ETSS water level forecast in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at twelve CO-OPS water level stations (Table A.1).

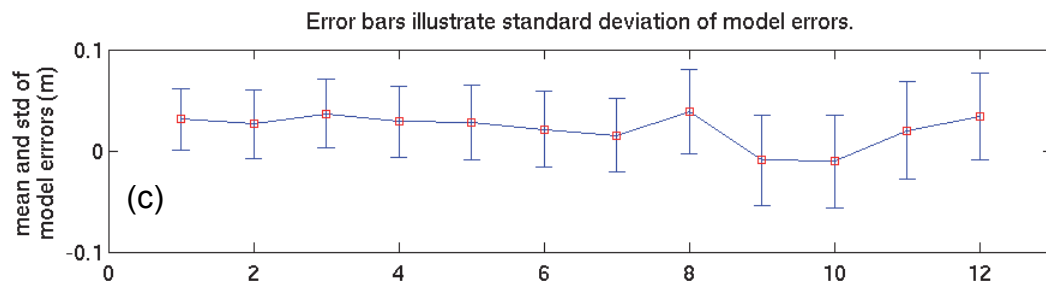
10/ 2012



01/2013



04/2013



07/2013

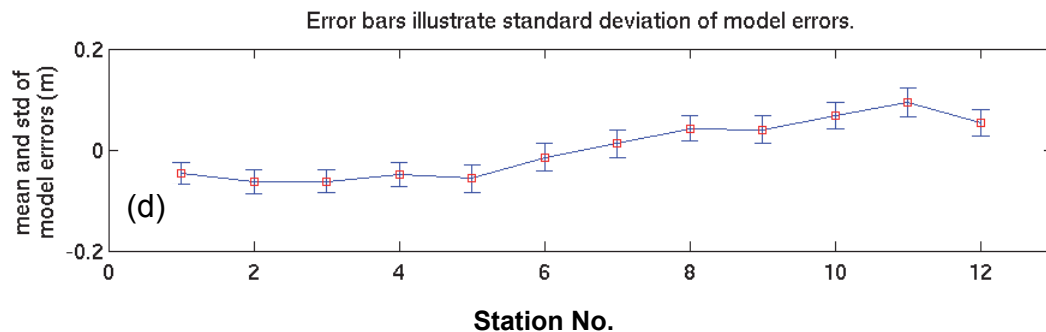


Figure 3.2. Means (red squares) and standard deviations (blue error bars) of the ETSS water level forecast errors in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 12 CO-OPS water level stations (Table A.1).

RMSE of Individual Stations In the previous passage, we discussed the model skill averaged over all stations. In the following, we examine the skill at each individual station. Figures 3.4 - 3.7 display station maps of water level RMSE for (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013. The four plots in each figure correspond to forecast hours 24, 48, 72, and 96.

In general, the model demonstrates similar skills at hour 24 across all stations and seasons with RMSE generally less than 6 cm.

At each station, model skill gradually degrades with increasing forecast hour. It demonstrates favorable skills within the first two days of a forecast cycle. For instance, for all four months, the RMSE ranges between 2 and 8 cm at both hours 24 and 48. The RMSE gradually increases with the increasing forecast hour; the RMSE ranges between 6 and 12 cm at hour 72 and reaches as high as 15 cm at some stations at hour 96.

However, the degradation varies by month. It was most severe in January and April, less severe in July, and at an immediate level in October. For example, we examined the case at Station 9418767 in Humboldt Bay, California. This station demonstrates intermediate magnitudes of RMSE and is thus quite representative. At this station, from hours 24 to 96, the RMSE dramatically increases from 4 to 10 cm in January and April; however, the RMSE increases much more gradually, from 3 to 6 cm, in July and October.

The model performance varies by latitude. It demonstrates similar skill at the ten southern stations located between 35 °N and 45 °N (see Table A.1 for the station meta data) and much less satisfactory skill at the four northernmost stations, namely, stations 9410230, 9410840, 9412110, and 9413450. The spatial disparity is most noticeable for hours 72 and 96. For instance, at the four northern stations, the RMSE on forecast hour 24 throughout January 2013 (Figure 3.5) was about 7 cm, slightly greater than 4-5 cm at the remaining stations to the south. The RMSE increased to 10-15 cm at hours 72 and 96, as compared with 5-9 cm at the remaining stations.

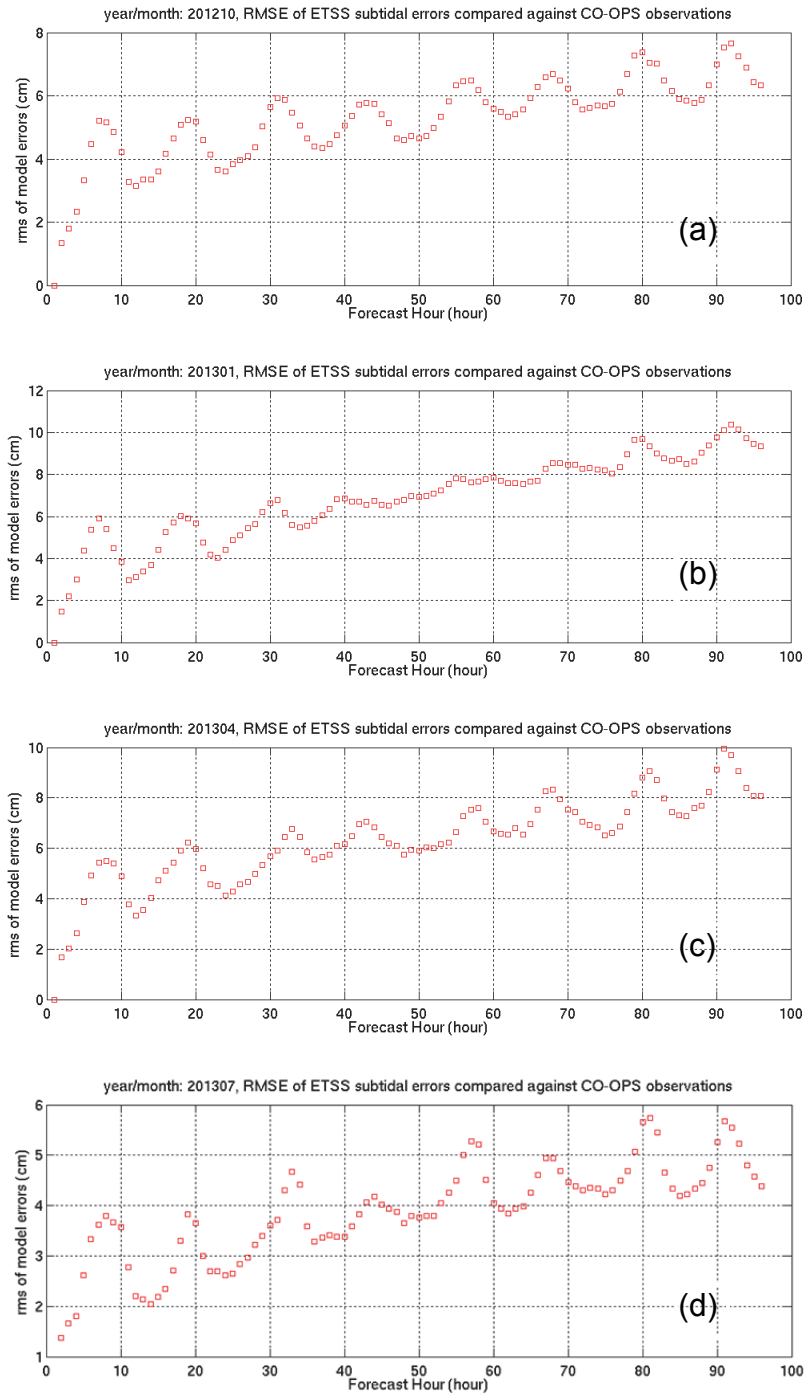


Figure 3.3. The RMSE of the ETSS water level forecast at each forecast hour in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013.

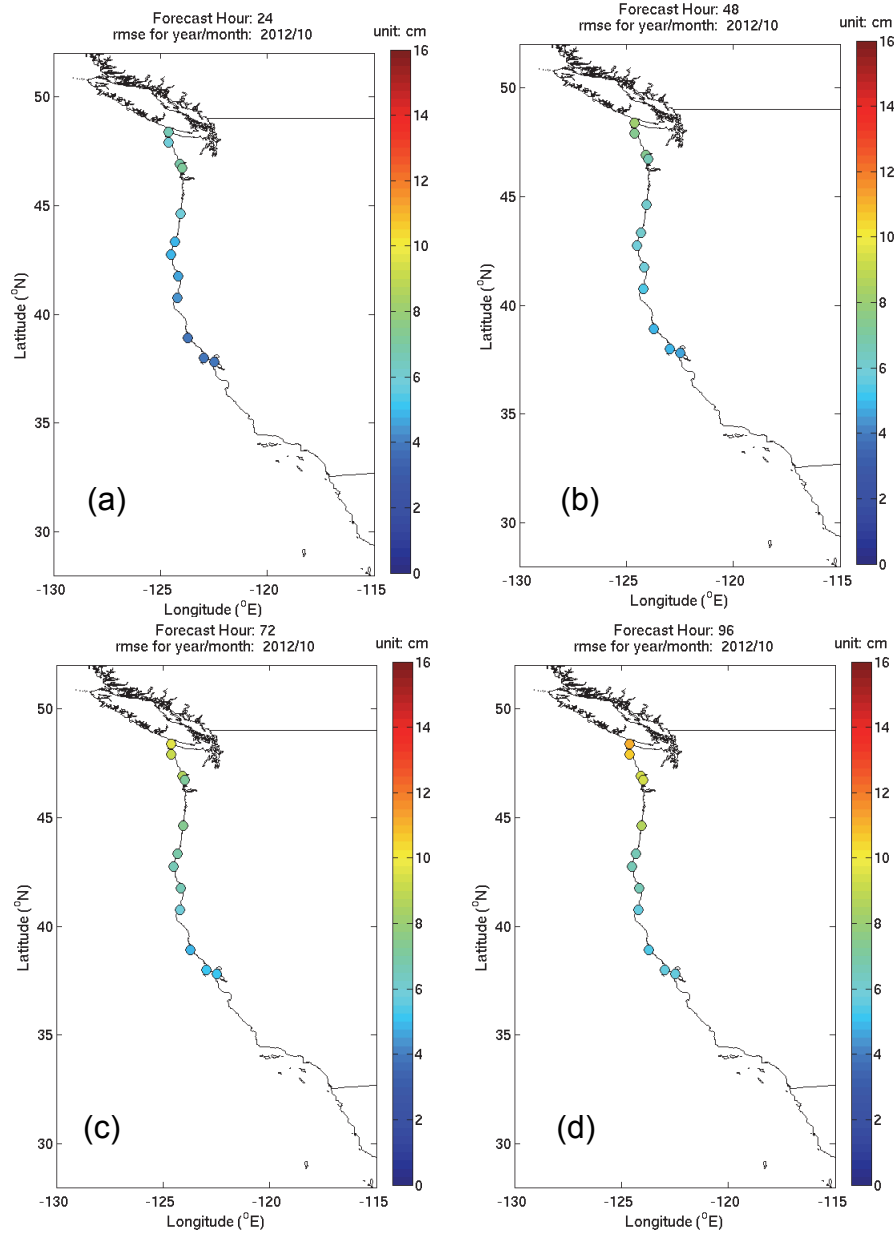


Figure 3.4. Color coded RMSE maps of the ETSS water level forecast in October 2012. The four plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, and (d) 96, respectively.

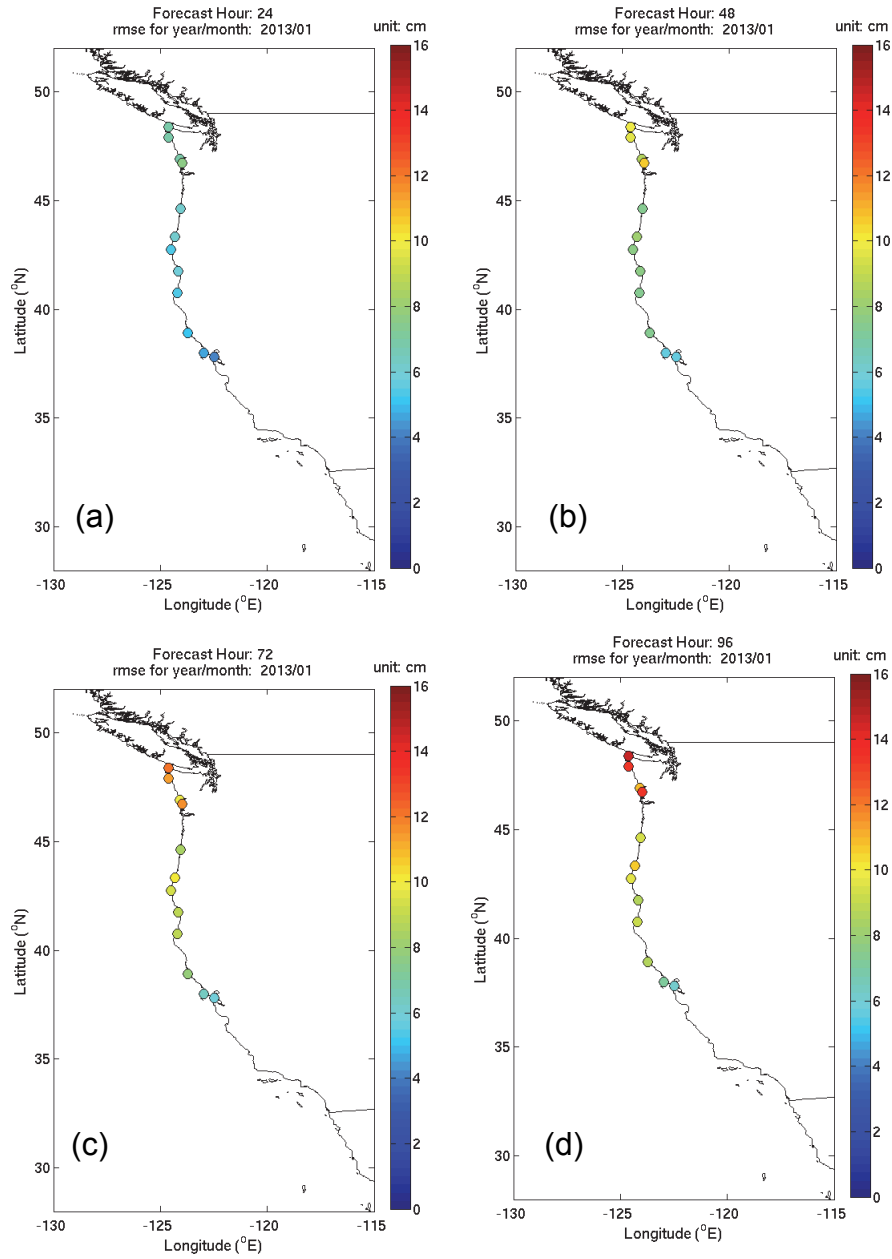


Figure 3.5. Color coded RMSE maps of the ETSS water level forecast in January 2013. The four plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, and (d) 96, respectively.

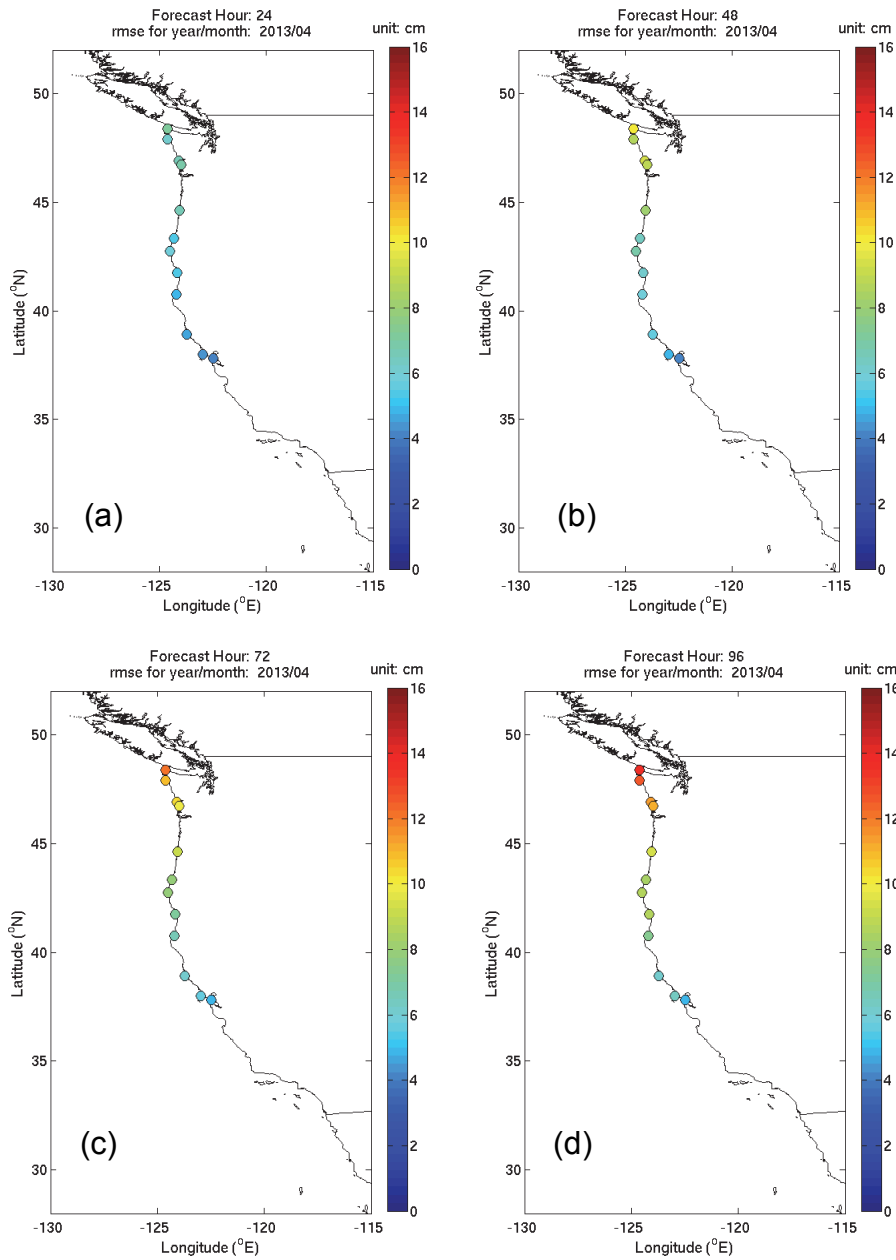


Figure 3.6. Color coded RMSE maps of the ETSS water level forecast in April 2013. The four plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, and (d) 96.

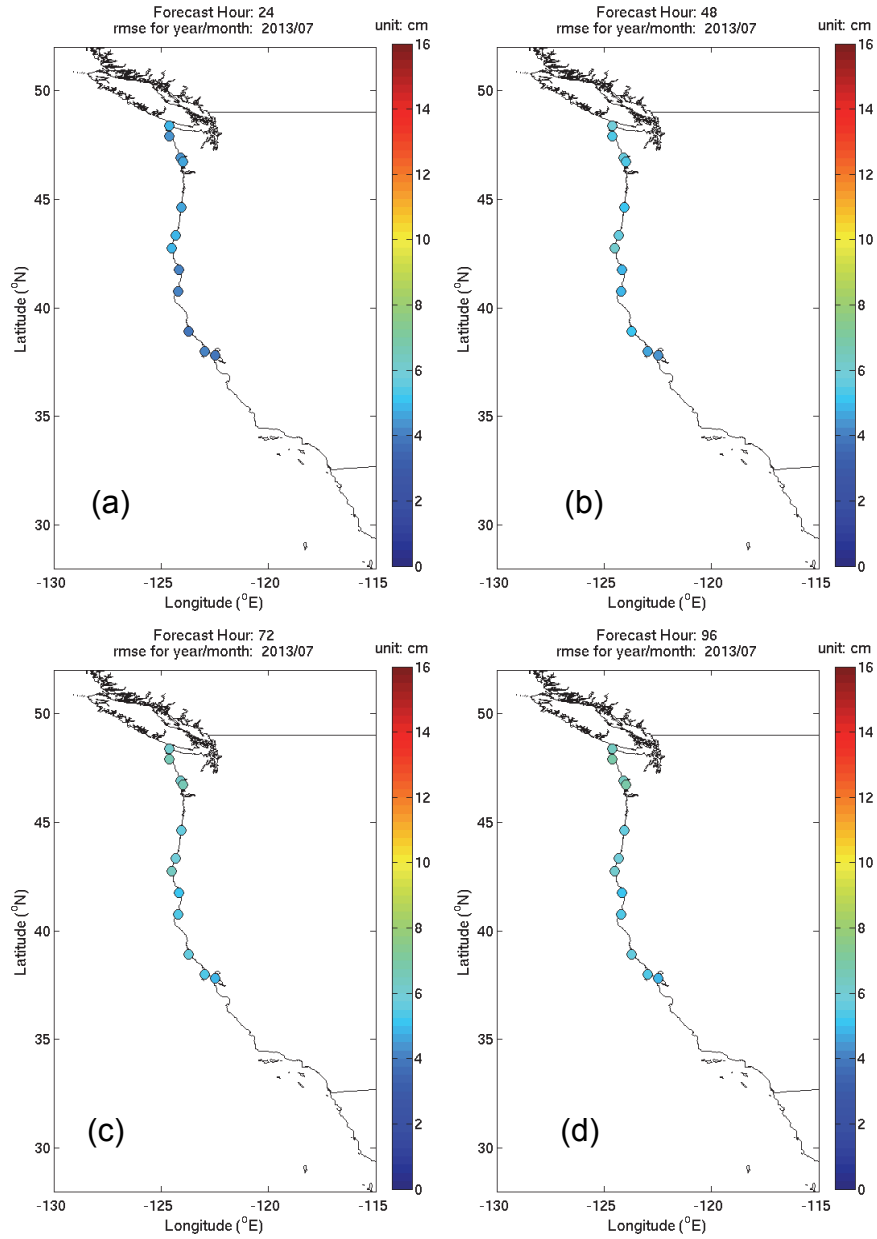


Figure 3.7. Color coded RMSE maps of the ETSS water level forecast in July 2013. The four plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, and (d) 96.

3.2. G-RTOFS Forecast

In this section, we compare the G-RTOFS water level forecasts with measurements at 18 CO-OPS water level stations (Table B.1). Table B.1 lists the station IDs, names, and geographical locations in longitude and latitude.

3.2.1. FC Based Assessment

Figures 3.8(a) - (d) display model RMSE maps in October 2012, January, April, and July 2013. The RMSE typically ranges between 6-15 cm in October and July and between 10 and 20 cm in January and April. This demonstrates a better G-RTOFS performance in fall and summer than in winter and spring.

The maximum RMSE in each season is 16 cm, 19 cm, 20 cm, and 15 cm in October, January, April, and July, respectively. In addition, RMSE also demonstrates noticeable spatial disparity. It varies in a 9-cm range in July and October and in a 10-cm range in January and July.

It is worth noting that the model skill appears to be least satisfactory at the four northernmost stations in April and July. The corresponding RMSE ranged between 12 cm and 20 cm. We attribute this anomalous model-data discrepancy to the ocean dynamics near the Strait of Juan de Fuca being not properly resolved by G-RTOFS.

Figures 3.9 (a) - (d) display the mean and STD of model errors at the 18 CO-OPS water level stations in October 2012, January, April, and July 2013. The mean model error ranges between 0 and 18 cm indicating that G-RTOFS overpredicted subtidal water levels throughout all four seasons.

The STD ranges from around 2 to 10 cm. In each of the four months it increased from the southern to the northern stations. This trend is more evident in January and April than it is in October and July. For instance, in April 2013, the STD gradually increased from 3 cm at the southernmost station (Station 1, ID: 9410230) to an intermediate value of 6 cm at a mid-latitude station (Station 189, ID: 9416841) to 11 cm at the northernmost station (Station 18, ID: 9440910).

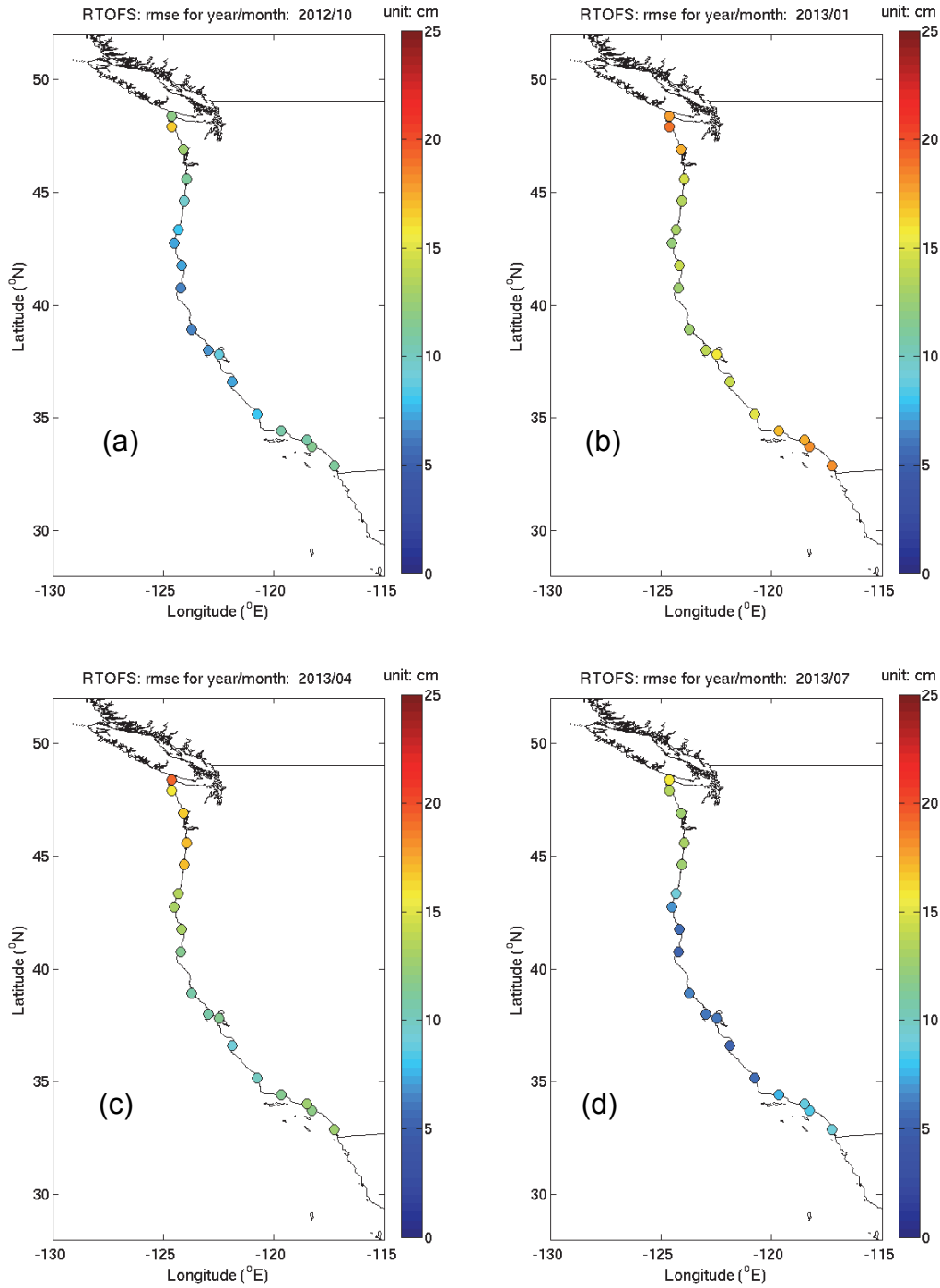
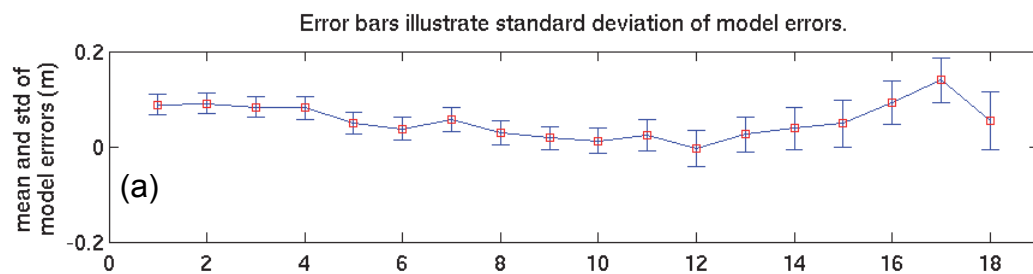
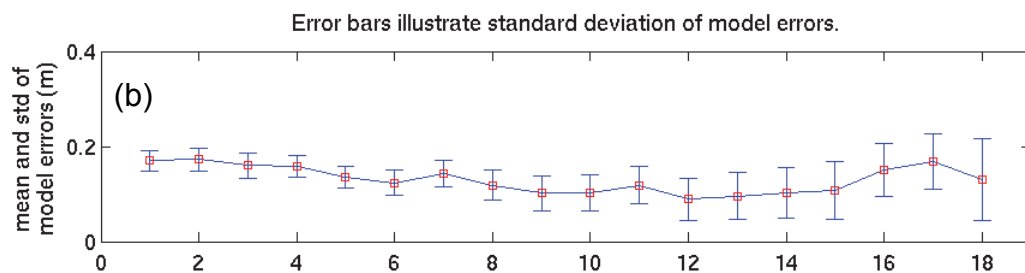


Figure 3.8. Color coded RMSE maps of G-RTOFS water level forecasts in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 18 CO-OPS water level stations (Table B.1).

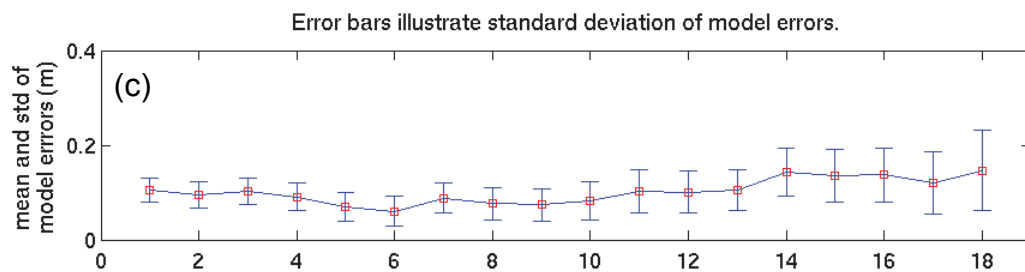
10/2010



01/2013



04/2013



07/2013

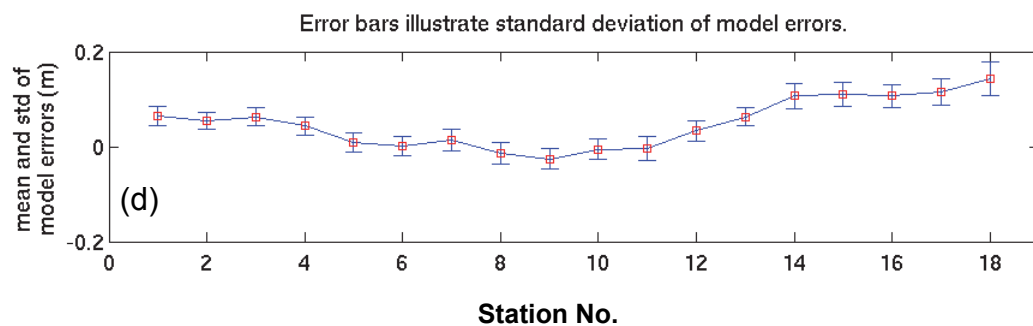


Figure 3.9. Mean (red squares) and standard deviation (blue error bars) of the G-RTOFS water level errors for 2013 at 18 CO-OPS water level stations from: (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013.

3.2.2. FH Based Assessment

In this section, we apply the FH method to investigate the G-RTOFS performance for SSH at individual hours within the forecast cycle.

Station-Averaged RMSE Figure 3.10 displays station-averaged RMSE at each forecast hour in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013. In all four months, the RMSE first increases monotonically as the forecast hour extends and then, upon reaching a transition point, it remains invariant for the remainder of the forecast cycle. This means that the model skill degrades with increasing forecast hour until the RMSE reaches a limiting value. The specific level of plateau and the transition hour vary by month.

The ranking of months with regard to G-RTOFS RMSE (limiting value of RMSE) performance is as follows: July, October, January, and then April. The corresponding RMSE was around 4.5 cm, 6 cm, 10 cm, and 10 cm, respectively. The number of hours taken to reach the limiting value was around 60, 60, 100, and 124, respectively.

RMSE of Individual Stations In the above, we discussed station-averaged model performance. The calculations reflect the overall model skill across the entire U.S. west coast region. In the following, we assess the model skill at the individual stations. Figures 3.11-3.14 display station maps of water level RMSE for (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013. Each figure consists of six plots which correspond to forecast hours 24, 48, 72, 96, 120, and 144.

In general, model results demonstrate favorable agreement with observations within the first two days of a forecast cycle. For instance, in all four months, typical RMSE magnitudes were less than 6 cm at hour 24 and less than 8 cm at hour 48. From there, the model performance gradually deteriorates with increasing forecast hour.

As the forecast hour increases, the skill decreases. However, the degree of degradation varies by month. Similar to the ETSS forecast (Section 3.1), the G-RTOFS performance was least favorable in January and April, most favorable in July, and average in October. For example, we examined the model performance at Station 9418767 which is located in Humboldt Bay, California. Amongst the 18 stations, this station demonstrates an intermediate range of model skill with regard to magnitude of RMSE. From hours 24 to 144, the RMSE increases dramatically from 5 cm to 11 cm in January and April. However, it increases more gradually from 2 to 6 cm in July and October.

The G-RTOFS forecast did not perform evenly among the stations. In general, it showed better skill at the 9 stations south of 40 °N than at the remaining stations to the north. Looking at the 9 northern stations, the model skill dropped from south to north. For instance, the RMSE map at hour 72 for April 2013 is shown in Figure 3.13(c). RMSE increased monotonically from 6 cm at the station near 40 °N to over 15 cm at the northernmost station near 48 °N.

It is worth noting that the model skill demonstrates the least spatial and temporal variability in the summer season; see Figure 3.14 for the July 2012 result. RMSE varied between 4 cm and 6

cm between forecast hours 24 and 144. Similarly it remained nearly invariant, $\sim 4\text{-}6$ cm, across all 18 stations.

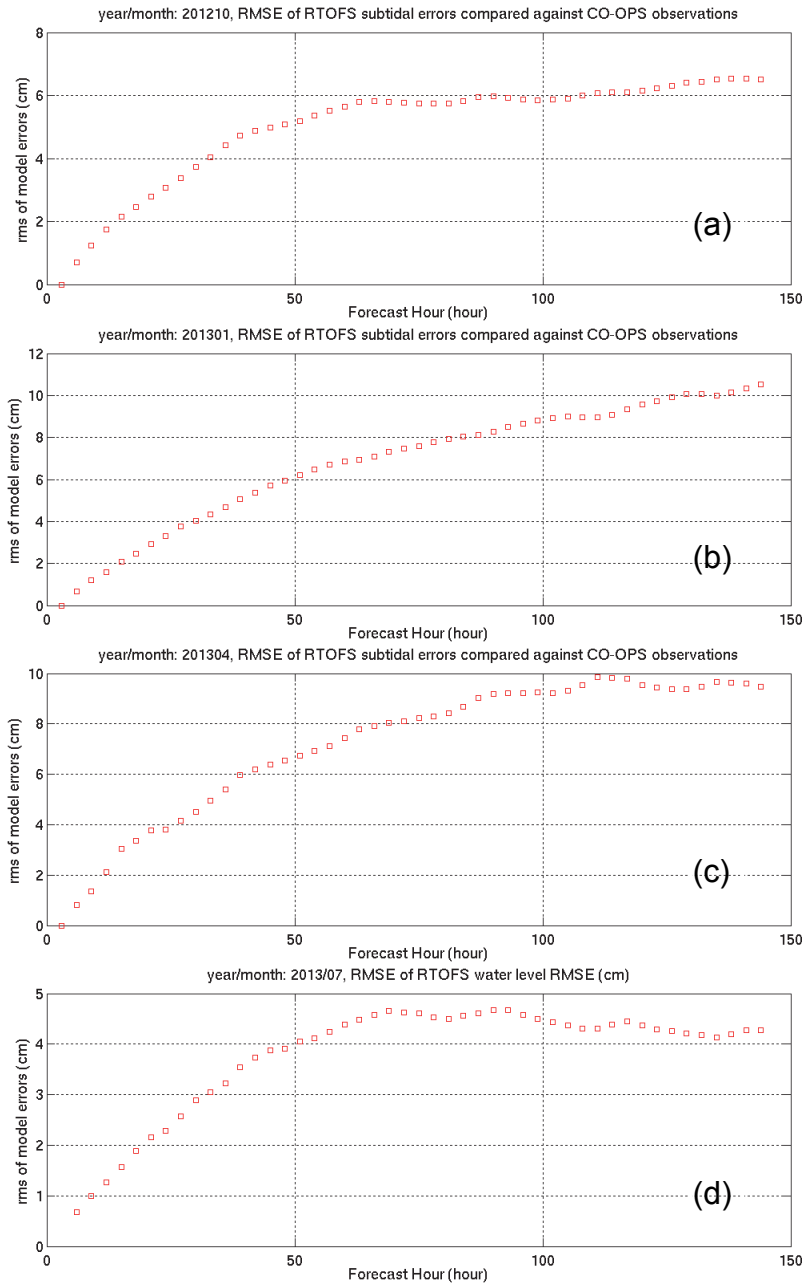


Figure 3.10. Station-averaged root mean squared errors of G-RTOFS water level forecasts in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013, respectively.

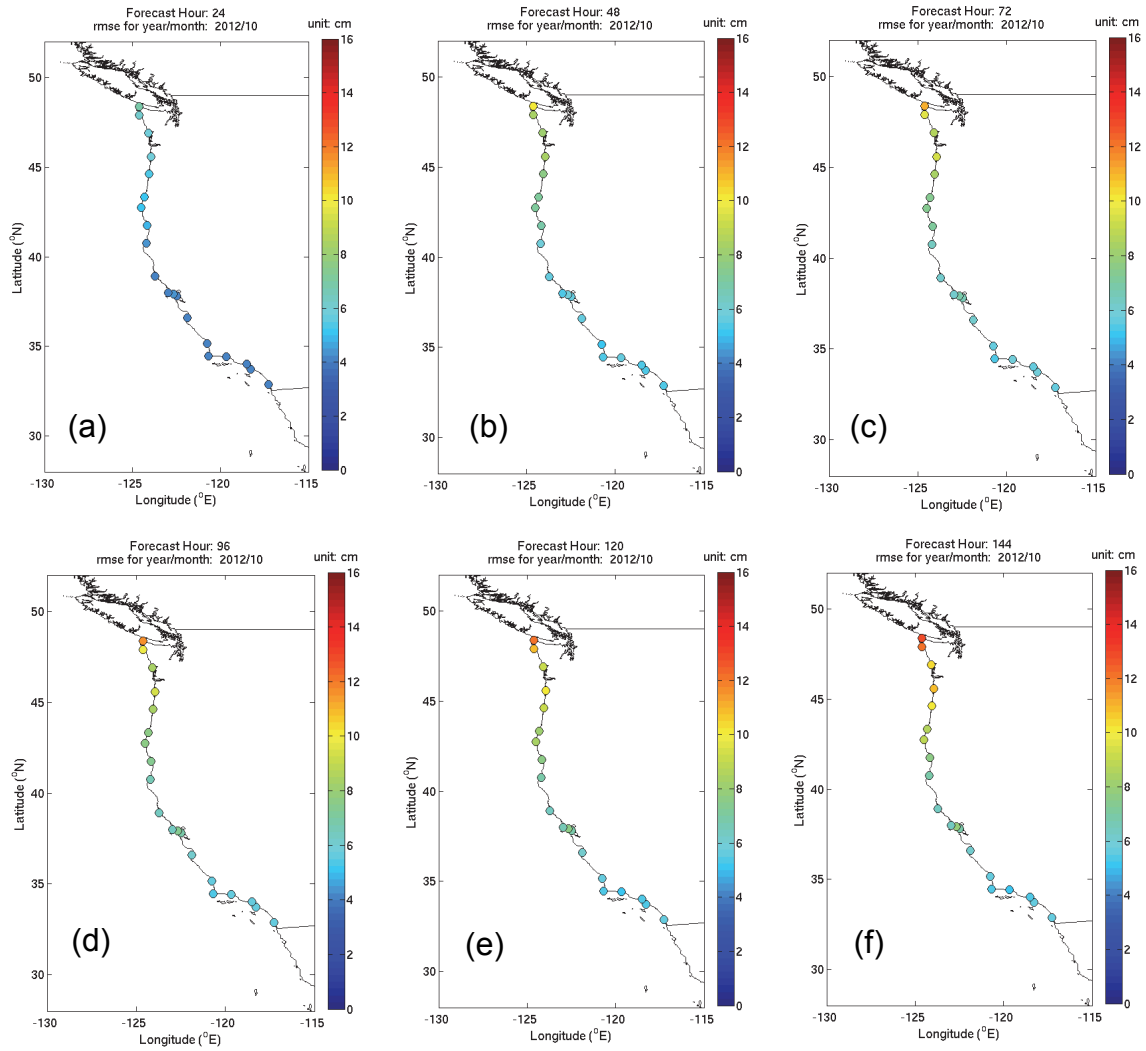


Figure 3.11. Color coded RMSE maps of G-RTOFS water level forecasts in October 2012. The six plots correspond to forecast guidance hours (a) 24, (b) 48, (c) 72, (d) 96 (e) 120 and (f) 144, respectively.

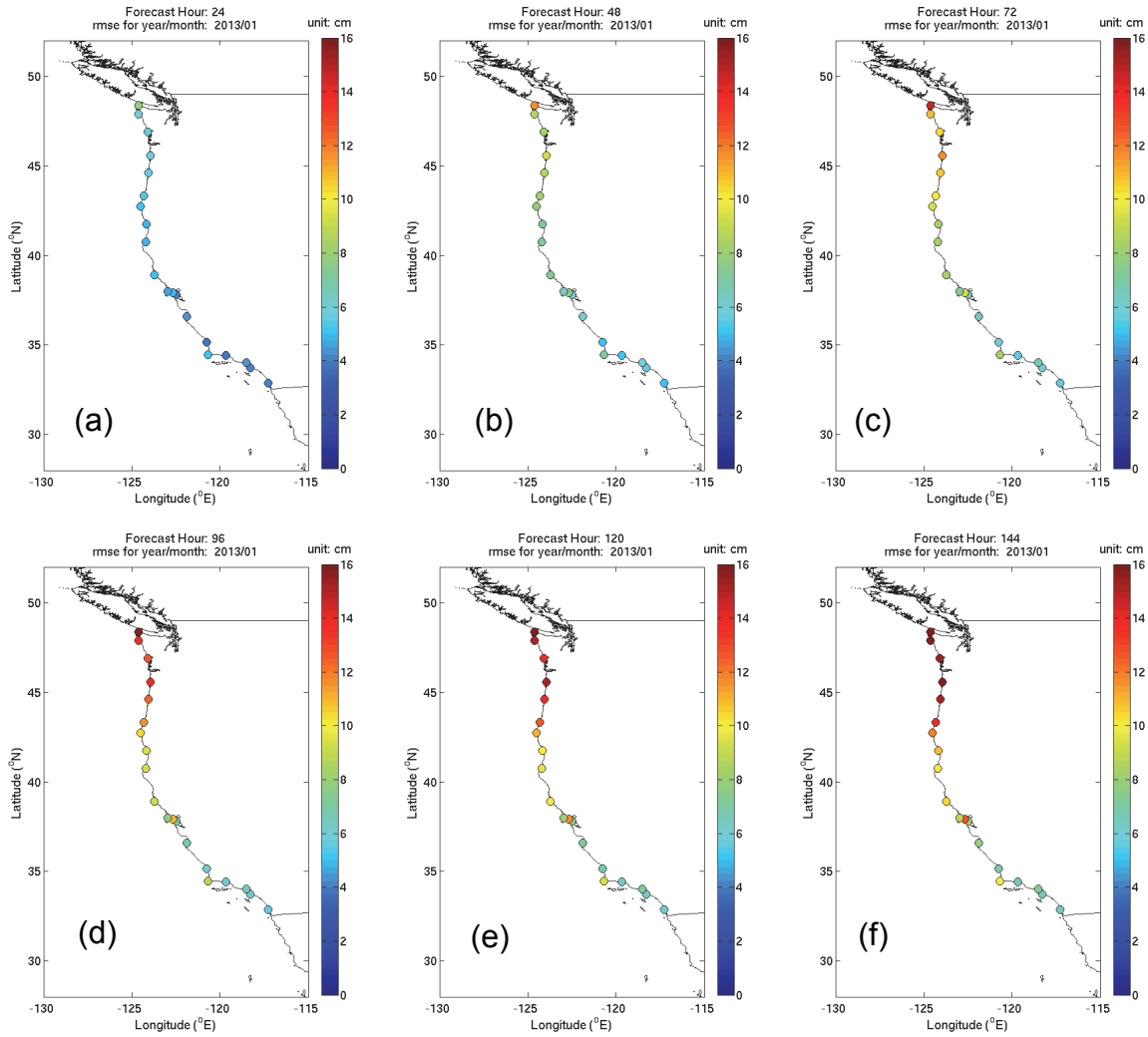


Figure 3.12. Color coded RMSE maps of G-RTOFS water level forecasts in January 2013. The six plots correspond to forecast guidance hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144, respectively.

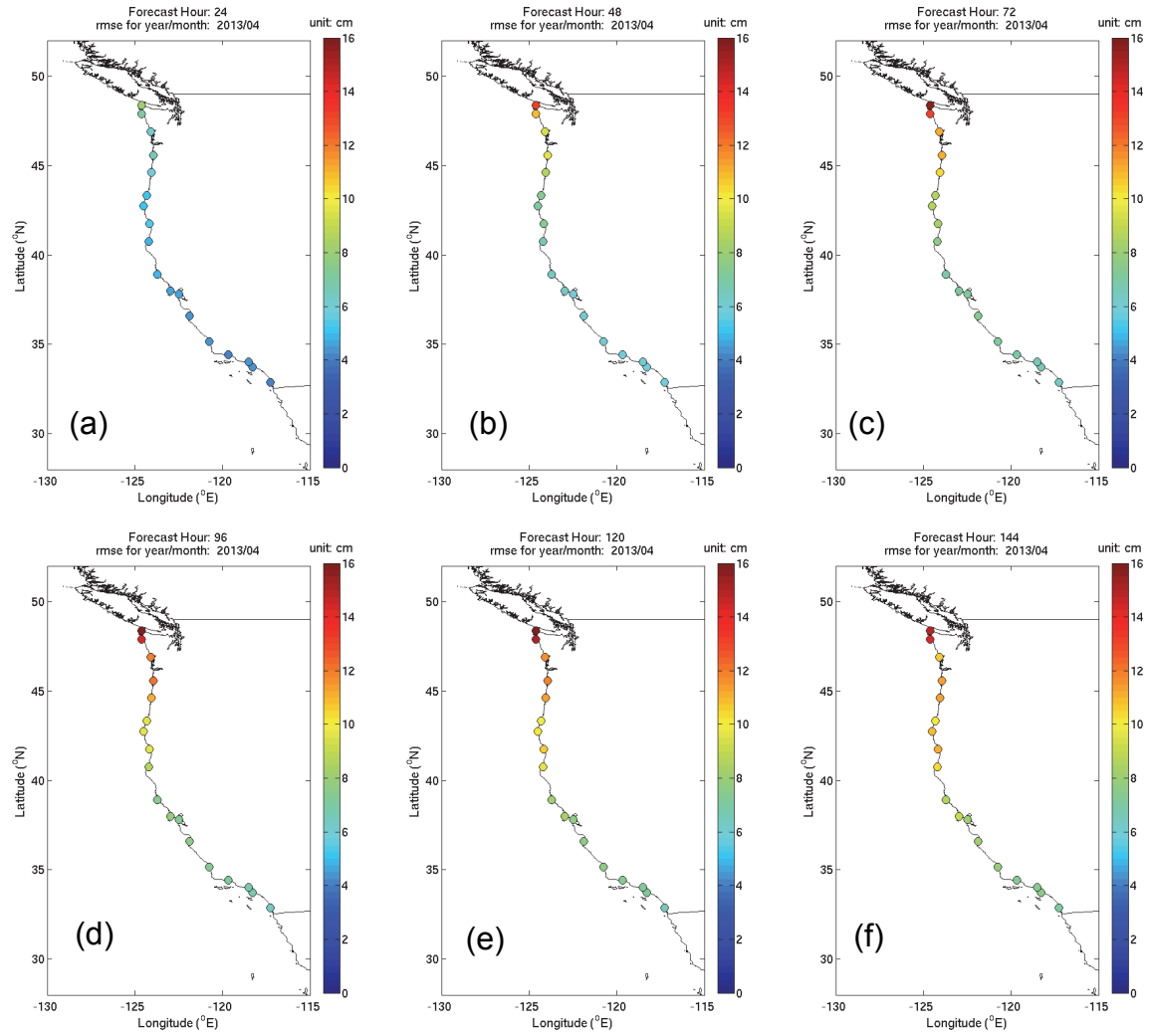


Figure 3.13. Color coded RMSE maps of G-RTOFS water level forecasts in April 2013. The six plots correspond to forecast guidance hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144, respectively.

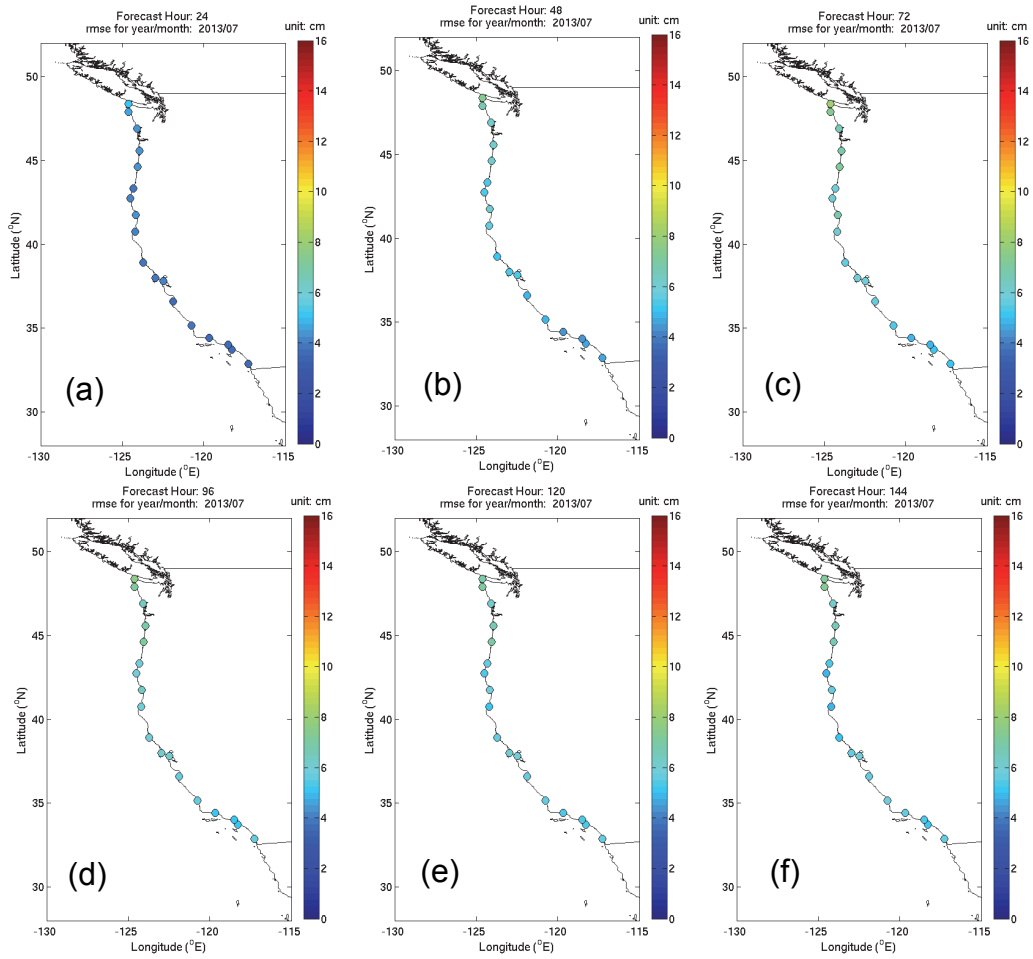


Figure 3.14. Color coded RMSE maps of G-RTOFS water level forecasts in July 2013. The six plots correspond to forecast guidance hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144, respectively.

3.3. Summary

We assessed the performance of the ETSS and the G-RTOFS SSH forecast guidance by comparing model results with observations at 12 (Table A.1) or 18 (Table B.1) CO-OPS water level stations. We applied both the forecast-cycle (FC) based method and forecast-hour (FH) based method in the analysis (Section 2).

Table 3.1 lists the station averaged bias, standard deviation, and root-mean-squared error for the fall, winter, spring, and summer estimated using the FC method. The bias of ETSS displays noticeable season variability; it is about -7 cm in the fall and winter and is 2.2 and 0.1 cm in spring and summer, respectively. In contrast, G-RTOFS displays a bias of around 5-13 cm through all four seasons. The STD of both models ranges between 2 and 4 cm through the four seasons. The RMSE for ETSS ($RMSE_{ETSS}$) increases from 6 to 7 cm in spring and summer to around 9 cm in the fall and winter. The RMSE for G-RTOFS ($RMSE_{G-RTOFS}$) increases from ~7 cm in fall and summer to 11-14 cm in the winter and spring.

Table 3.2 lists the RMSE of the four seasons estimated using the forecast-hour (FH) based method. Note that the FH based method focuses on investigating the evolution of model performance with the increasing hours of a forecast cycle. G-RTOFS and ETSS exhibit a similar level of performance skill. In general, the skill degrades as a forecast projects further into the forecast cycle. The $RMSE_{ETSS}$ increases from less than 2 cm at the beginning of a forecast cycle to 6-10 cm near the end of the cycle. Model skill exhibits seasonal variability. The maximum $RMSE_{ETSS}$ is 8 cm, 10 cm, 10 cm, and 6 cm in fall, winter, spring, and summer, respectively.

The $RMSE_{G-RTOFS}$ for G-RTOFS monotonically increases from less than 2 cm at the beginning of a forecast cycle. The $RMSE_{G-RTOFS}$ then starts to level off and plateaus from the mid-cycle until the end. Model skill exhibits seasonal variability. The maximum $RMSE_{G-RTOFS}$ is 6.5 cm, 11.5 cm, 10.0 cm, and 4.7 cm in fall, winter, spring, and summer, respectively.

Table 3.1 Forecast cycle-based statistics of the ETSS and G-RTOFS SSH errors.

	SSH of ETSS			SSH of G-RTOFS		
	bias (cm)	std (cm)	rmse (cm)	bias (cm)	std (cm)	rmse (cm)
Fall (Oct. 2012)	-6.8	3.3	9.0	5.3	3.3	7.3
Winter (Jan. 2013)	-7.0	4.2	9.1	13.1	4.0	14.0
Spring (April 2013)	2.2	3.9	6.7	10.2	4.2	11.6
Summer (July 2013)	0.1	2.6	5.9	4.9	2.2	6.4

Table 3.2 Forecast hour-based RMSE of the ETSS and G-RTOFS SSH.

	RMSE _{ETSS} (cm)	RMSE _{G-RTOFS} (cm)
Fall (Oct. 2012)	1.5-8.0	1.3-6.5
Winter (Jan. 2013)	1.8-10.0	1.3-11.5
Spring (April 2013)	1.8-10.0	1.0-10.0
Summer (July 2013)	1.3-6.0	0.7-4.7

Note: RMSE_{G-RTOFS} and RMSE_{ETSS} denote root-mean-squared errors of G-RTOFS SSH and ETSS SSH, respectively.

4. PERFORMANCE OF G-RTOFS FOR SEA-SURFACE TEMPERATURE

In this chapter, we discuss the performance of the G-RTOFS sea-surface temperature (SST) forecast.

4.1. Compared with CO-OPS Observations

We compared the G-RTOFS SST forecast with measurements from 13 CO-OPS meteorological observation stations (Figure 1.6). Table C.1 lists the station meta data including IDs, names, and geographical locations in longitude and latitude. In the following, we describe the results derived from both the forecast cycle (FC) and forecast hour (FH) based methods.

4.1.1. Forecast Cycle Based Assessment

Figures 4.1(a) - (d) display RMSE maps for October 2012, January, April, and July of 2013. The RMSE displays much greater magnitude and spatial variability in summer than in the other months. It is less than 1.5 °C in October and January, it is under 2.5 °C in April, and it is as large as 4.5 °C in July. The RMSE in summer showed a greater spatial variation (~ 3 °C) than in the other months (~ 1 °C).

Figures 4.2(a) - (d) display the mean and standard deviation (STD) of model errors at the 12 stations in October 2012, January, April, and July 2013. Generally, the average model-data difference ranges between 0 and 4 °C. Its magnitude is about 1-2.5 °C in October, January, and April and slightly larger, ~ 3 -4 °C, in July. This indicates that G-RTOFS has better skill in fall, winter, and spring than in summer.

The corresponding STDs display similar seasonal variability, with typical values around 1 °C in October, January, and April and much greater magnitudes of ~ 2.5 °C in July. This suggests that the model error is less scattered from the monthly mean in fall, winter, and spring than in summer.

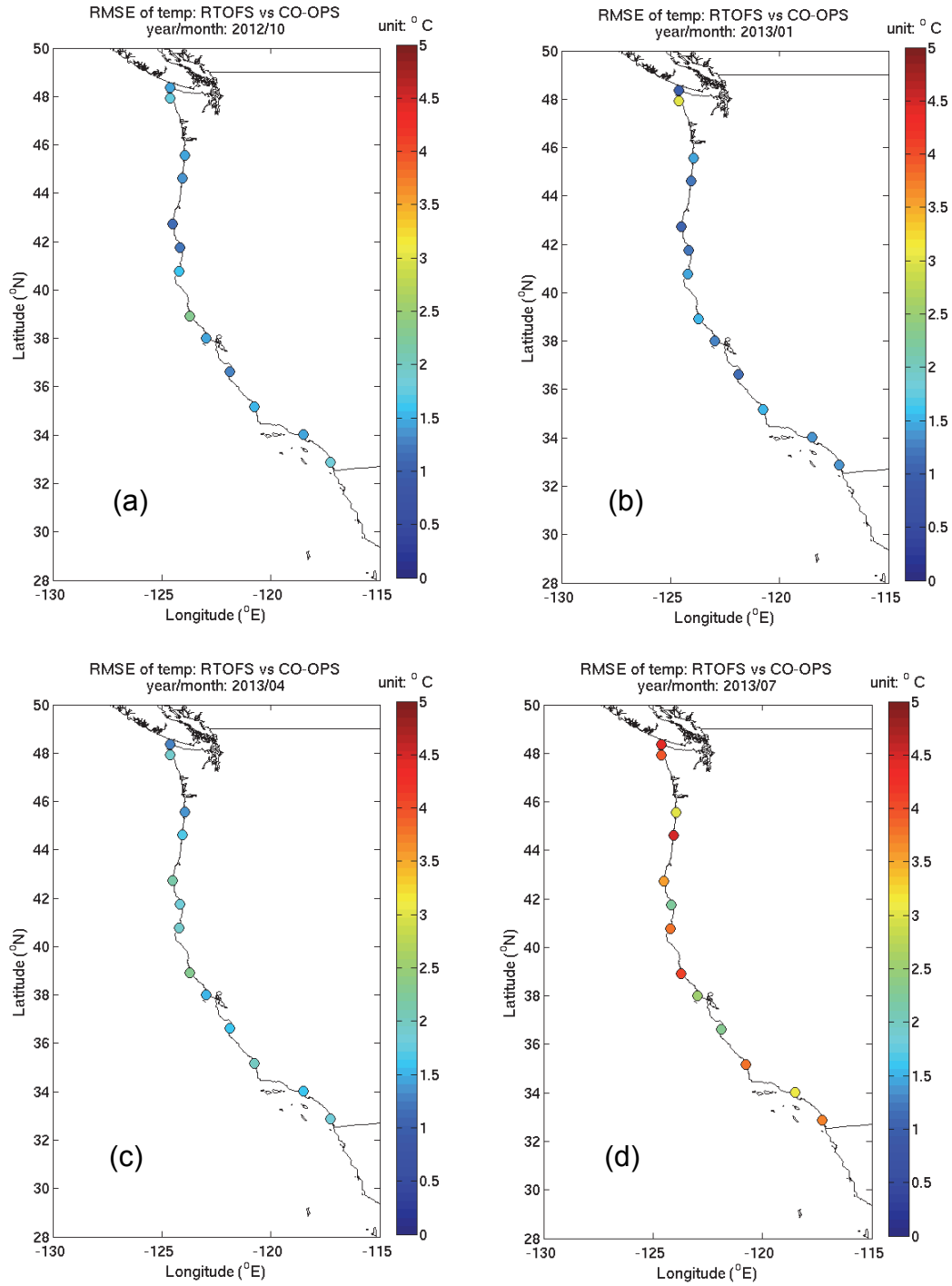


Figure 4.1. Color coded RMSE maps of the G-RTOFS SST forecast in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 13 CO-OPS meteorological observation stations (Table C.1).

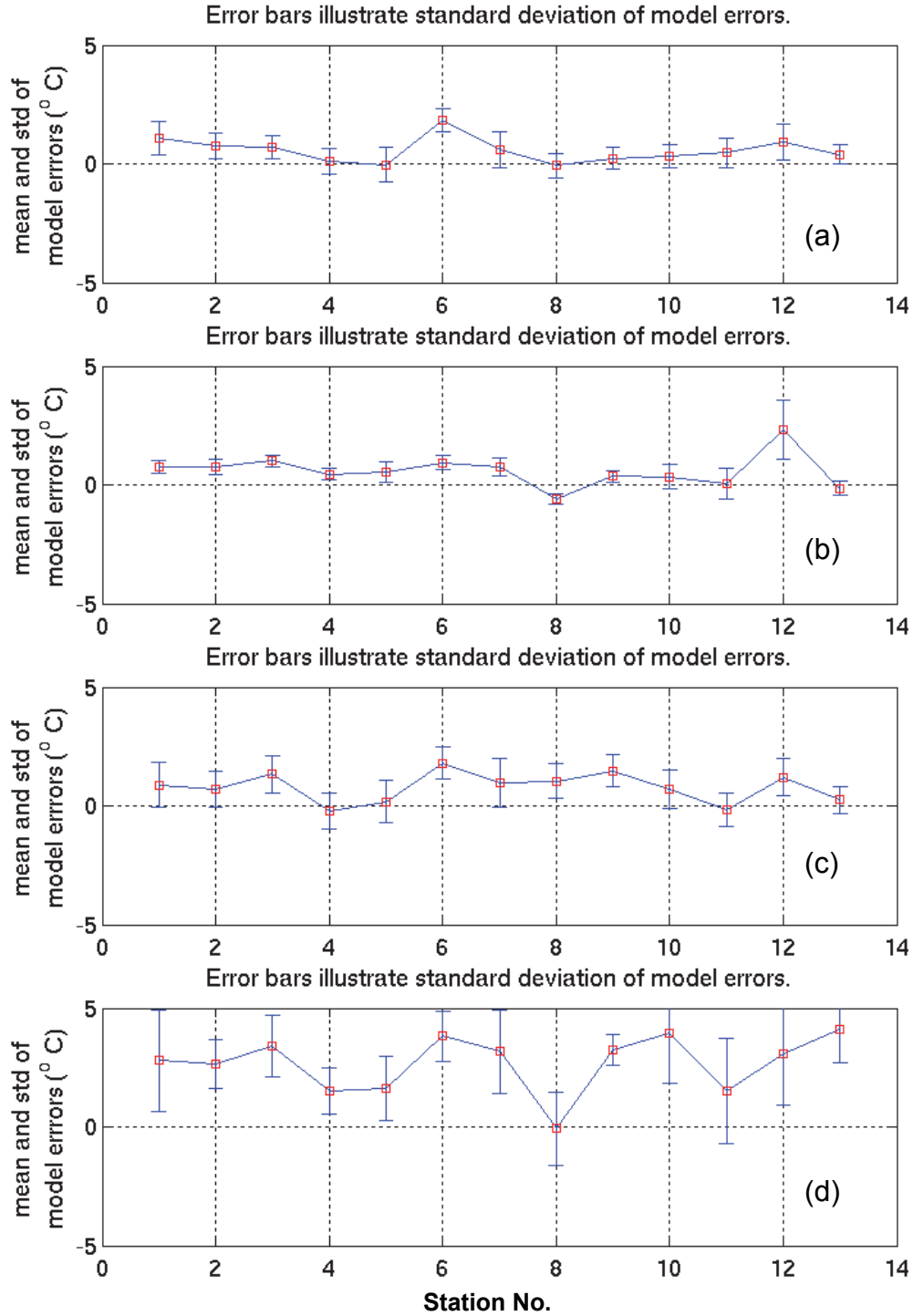


Figure 4.2. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SST errors in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 13 CO-OPS meteorological observation stations (Table C.1).

4.1.2. FH Based Assessment

In this section, we investigate the model performance at individual hours within a forecast cycle. We first describe the results in terms of station-averaged RMSE and then investigate results on a station by station basis.

Station-averaged RMSE Figure 4.3 displays station averaged RMSE at each forecast hour in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013. RMSE remains largely invariant (~ 1 - 1.5 °C) across the entire forecast cycle in October and January. It increases slightly, from ~ 1.5 °C at hour 3 to 2 °C at hour 144, in April. It exhibits significant variability, from ~ 2 °C at hour 3 to 4 °C at hour 144, in July.

Note that the RMSE displays a rather regular sinusoidal oscillation of the diurnal frequency. The oscillation amplitude varies by month; it is less than 0.2 °C in October and January, increases somewhat in April, and grows to as large as 1.5 °C in July. The oscillation does not reflect any inherent character of the model skill, but is a result of a numerical alias stemming from the low sampling rate (daily in this case) of the time series compiled for the present data analysis.

RMSE of Individual Stations In the above, we discussed the model skill averaged over all stations. In the following, we examine the skill at each individual station. Figures 4.4 - 4.7 display RMSE maps for all stations in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013. Six plots in each figure correspond to forecast hours 24, 48, 72, 96, 120, and 144, respectively.

For each of the four months, the RMSE remained nearly constant across all stations through the entire forecast cycle, even though the magnitude of RMSE increased significantly in July 2013. From hours 24 to 144, RMSE ranged between 1 and 2.5 °C in October and January, 1.5 and 2.5 °C in April, and between 2.0 and 5.0 °C in July. The model skill appears to be best in October and January and somewhat less satisfactory in July. Water stratification is less pronounced during fall and winter, and intensifies during spring and summer. We attribute the less favorable model performance in summer to its failure to resolve water stratification in coastal waters.

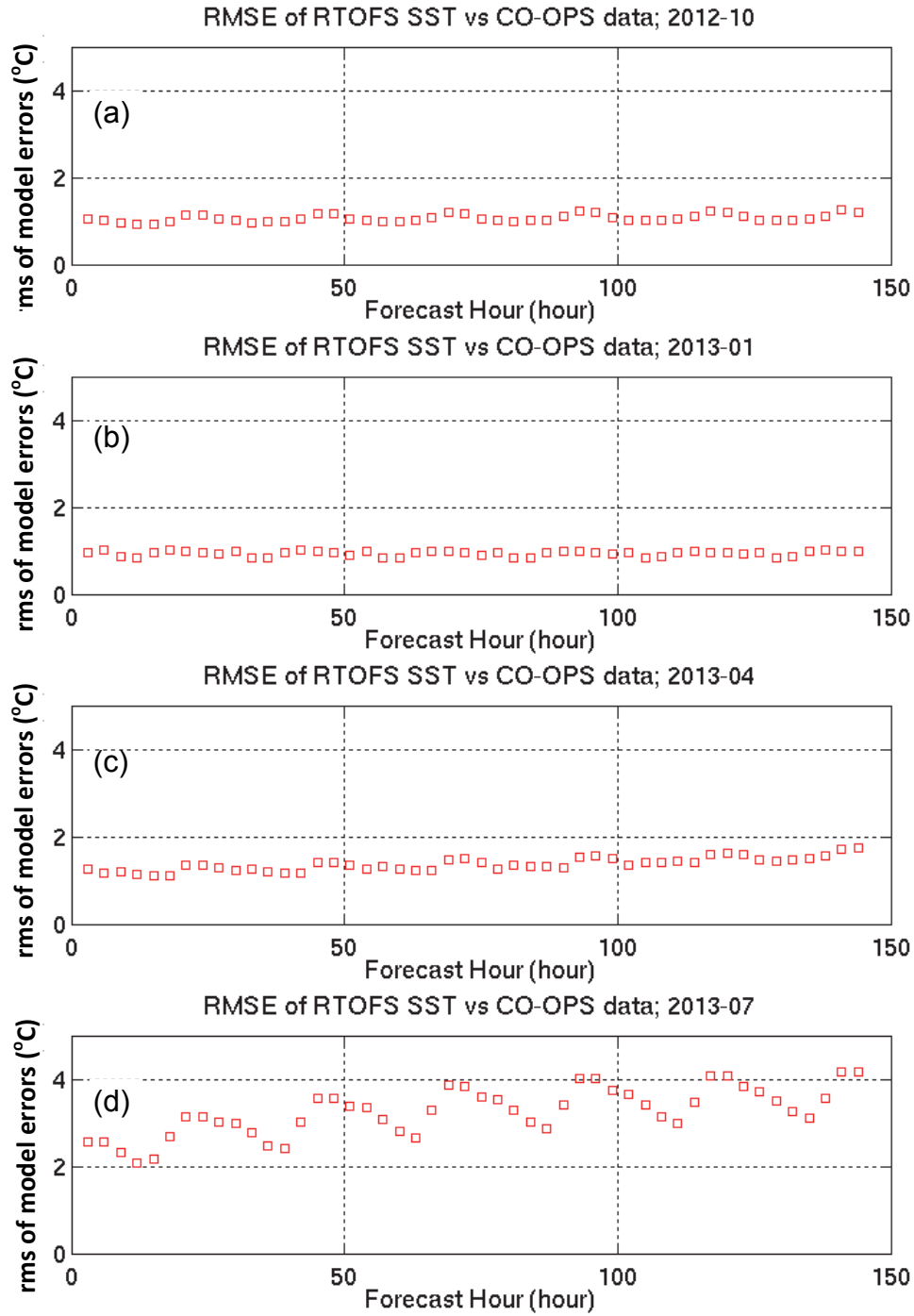


Figure 4.3. RMSE of the G-RTOFS SST forecast at forecast hours 1-144 in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013.

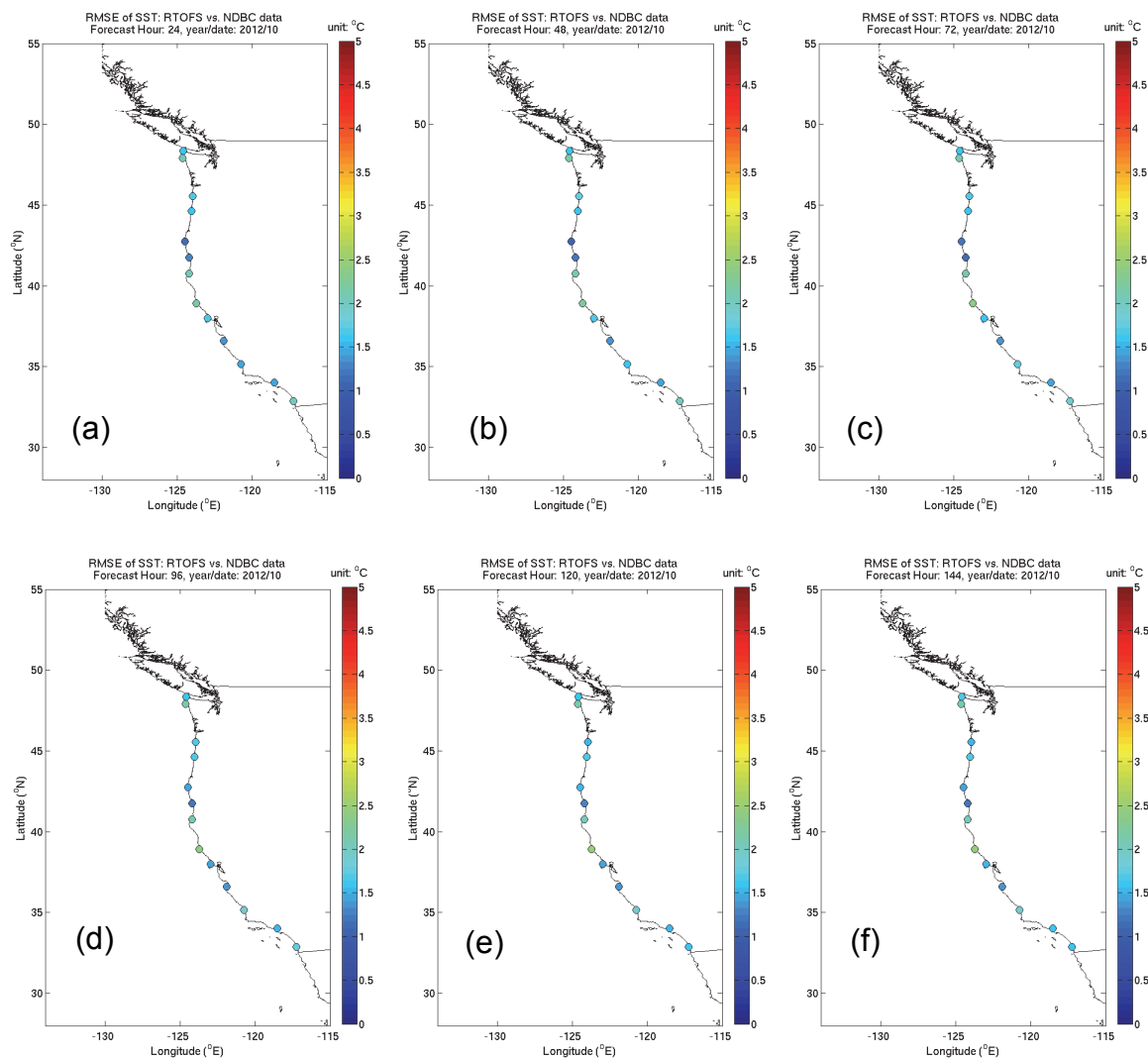


Figure 4.4. Color coded RMSE maps of the G-RTOFS SST forecast in October 2012. The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144.

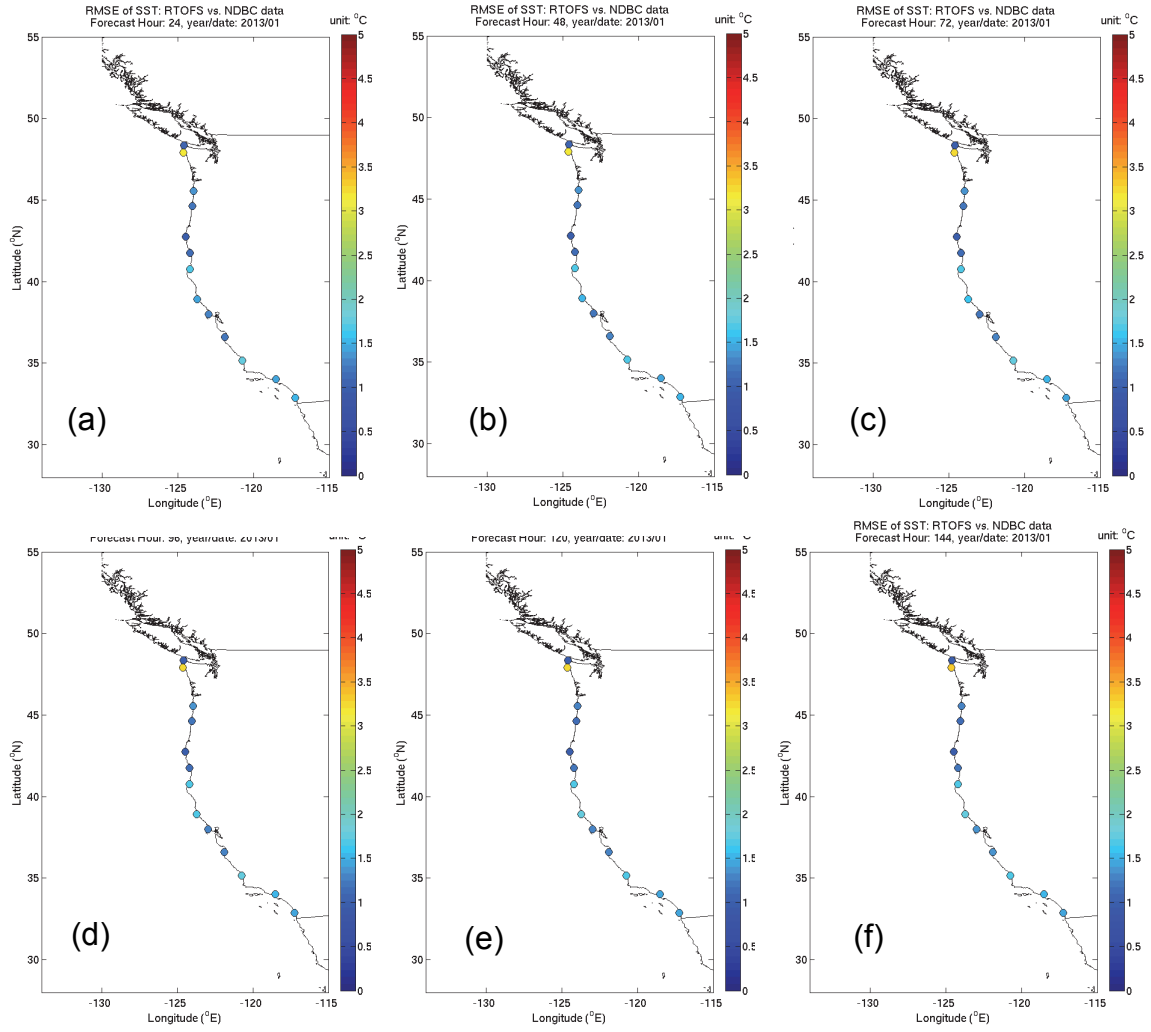


Figure 4.5. Color coded RMSE maps of the G-RTOFS SST forecast in January 2013. The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144.

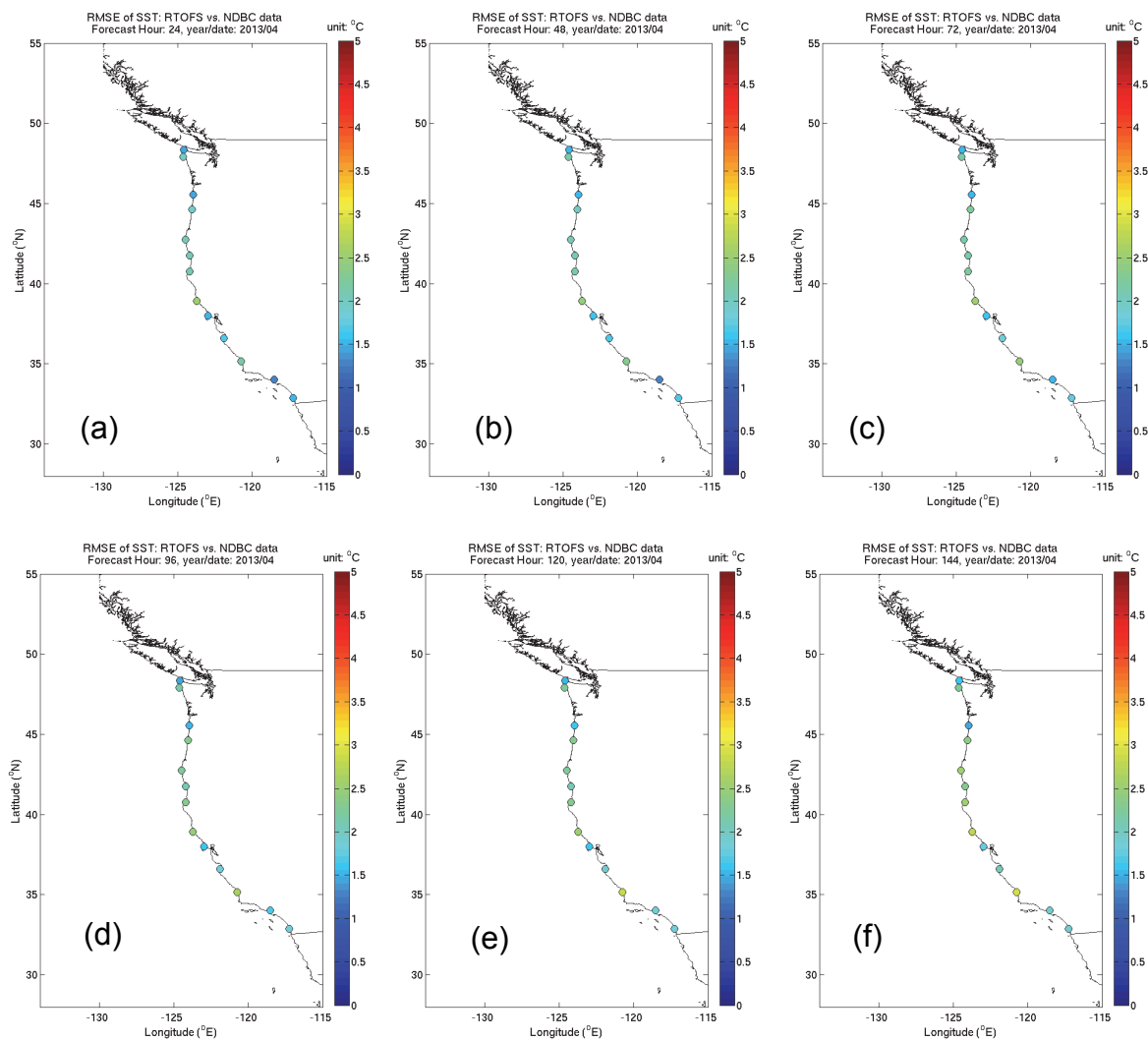


Figure 4.6. Color coded RMSE maps of the G-RTOFS SST forecast in April 2013. The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144.

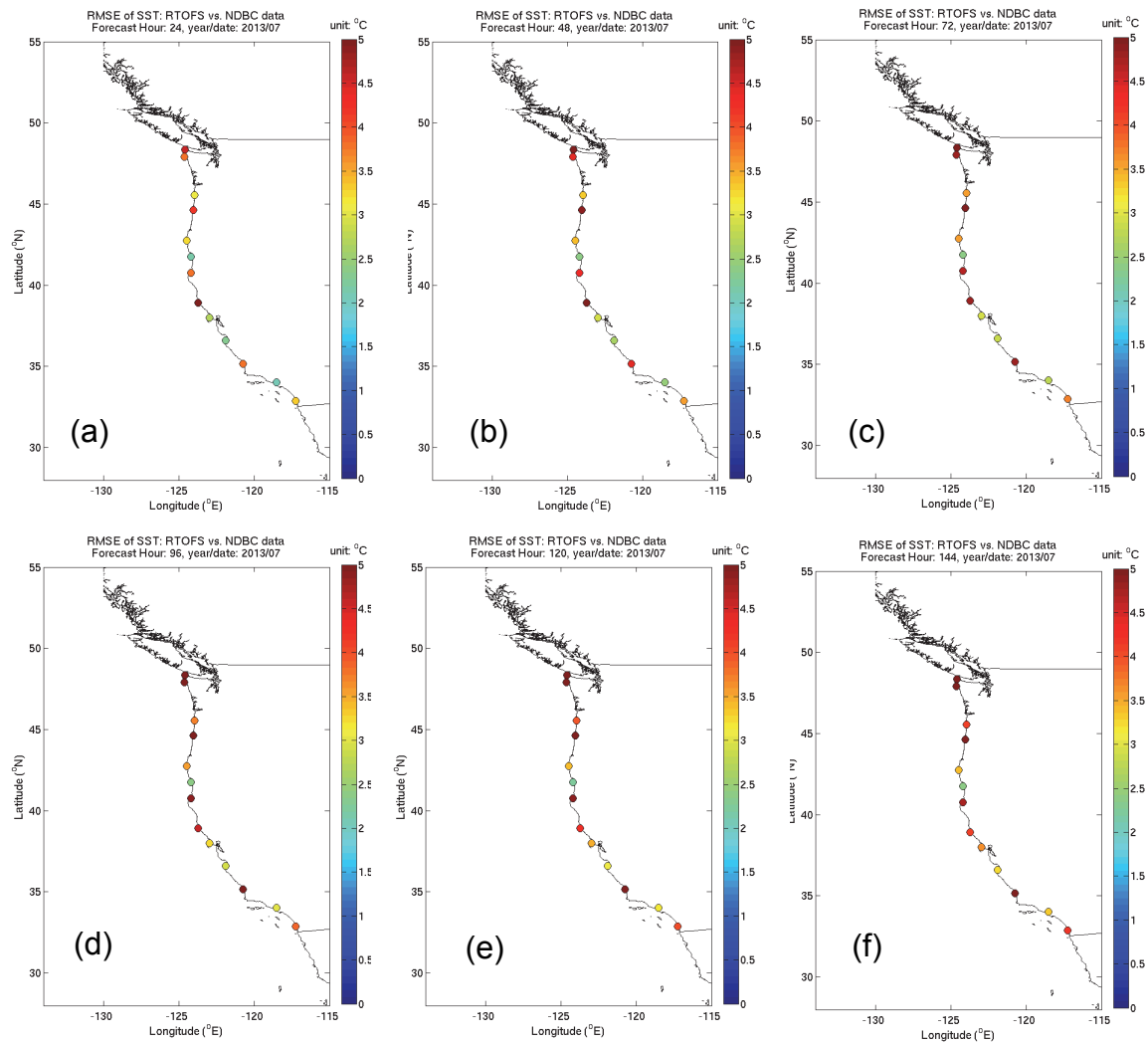


Figure 4.7. Color coded RMSE maps of the G-RTOFS SST forecast in July 2013. The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144.

4.2. Compared with NDBC Buoy Observations

We compared the G-RTOFS SST forecast with measurements at 34 NDBC buoys (Figure 1.7). Table D.1 lists the station IDs, names, and geographical locations in longitude and latitude. In the following, we describe the results using both the forecast cycle (FC) and the forecast hour (FH) based methods.

4.2.1. FC Based Assessment

Figures 4.8(a) - (d) display RMSE maps for October 2012, January, April, and July 2013. The average RMSE in summer (July) is much larger in amplitude and has greater spatial variability than in other months. The RMSE was typically less than 3 °C in October, January, and April and appeared as large as 4.5 °C in July.

The RMSE remained within a range of 1 to 2.5 °C across all stations in October and January. It displayed a two-mode spatial pattern in April, with RMSE at southern stations being about 1 °C greater than that of the northern stations. In July, the RMSE exhibited greater spatial variability than the other three months. It was around 1-2 °C at stations located north of 45°N, 2-3 °C at stations south of 35°N, and 3-5 °C in the intermediate latitudes.

Figures 4.9(a) - (d) show the mean and standard deviation of model errors at NDBC stations in October 2012, January, April, and July 2013. Table D.1 list IDs, names, and geographical locations in longitude and latitude of the 34 stations. The typical mean model-data difference ranged between -0.1 - 5 °C with some variation by month. The average model error was greater than zero, indicating that the model mean is biased higher than the mean SST of the observed data and that the model tended to overpredict SST.

The mean error ranged between -2 °C and 3 °C in October and January, between -1 °C and 2.5 °C in April, and appeared to be considerably greater, between 0 and 5 °C, in July. This indicates that G-RTOFS demonstrated better skill in fall, winter, and spring than in summer.

The corresponding STDs display similar seasonal variability with values typically less than 0.5°C in October, January, and April and greater magnitudes of ~1.5 °C in July. The model SST values were more closely confined to a statistical mean value in fall, winter, and spring than in summer.

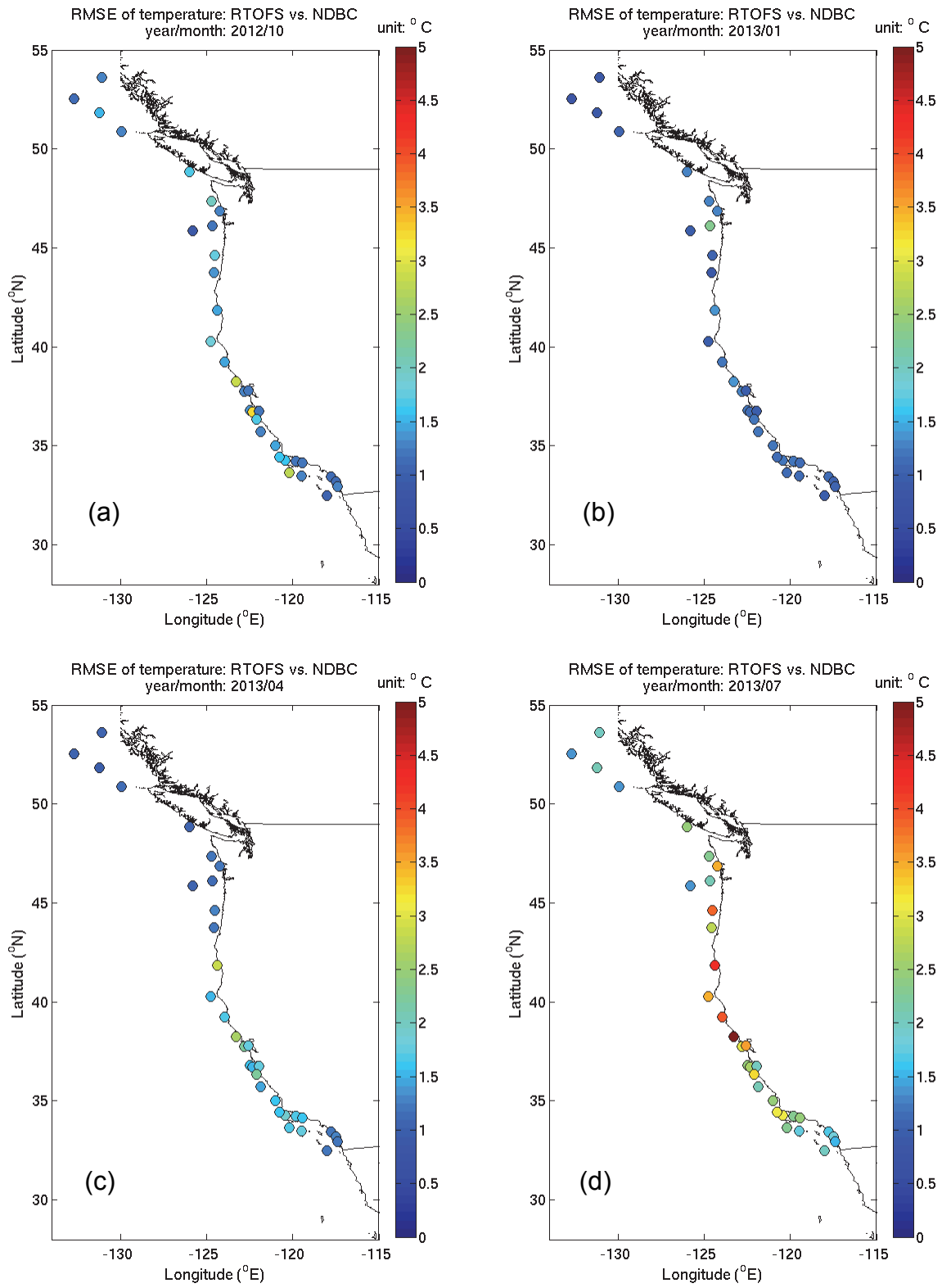


Figure 4.8. Color coded RMSE maps of the G-RTOFS SST forecast in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 34 NDBC stations (Table D.1).

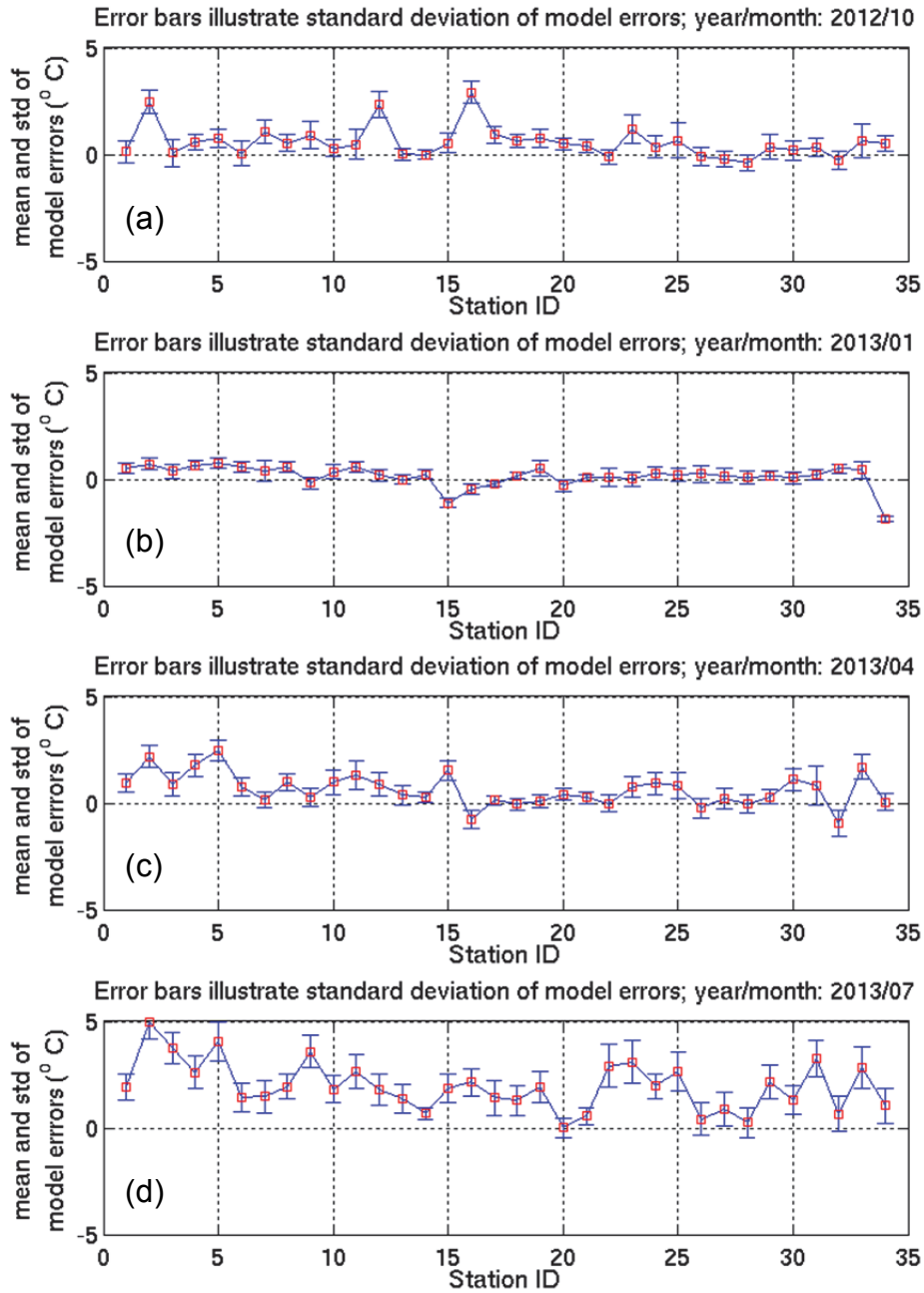


Figure 4.9. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SST errors in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 34 NDBC stations (Table D.1).

4.2.2. FH Based Assessment

In this section, we discuss the evolution of the model performance with increasing forecast hour. We first describe the results in terms of station-averaged RMSE and then investigate details at individual stations.

Station-averaged RMSE Figure 4.10 displays RMSE averaged over all 34 stations at each forecast hour in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013. RMSE remained largely invariant in both October and January: between 1 and 1.1 °C across the entire forecast cycle in October, and slightly smaller (~0.6 - 0.7 °C) in January. RMSE was larger in both April and July: between 0.8 and 1.5 °C in April, and between 1.8 and 2.7 °C in July.

In each forecast cycle, RMSE for all four months increases from hour 3 to hour 144. The model skill becomes less favorable with increasing forecast hour. This is most evident in the July result; RMSE increased from 1.8 °C at hour 3 to 2.7 °C at hour 144.

Note that the RMSE exhibits a rather regular sinusoidal perturbation of a diurnal frequency. The perturbation amplitude varied over different months. It was less than 0.1 °C in October and January, nearly 0.2 °C in April, and as large as 0.3 °C in July. We judged that the oscillations do not reflect any inherent character of the model skill, but instead, result from an alias due to the low sampling rate (which was daily in this case) of the underlying time series compiled for the present data analysis.

RMSE of Individual Stations In the above paragraphs, we discussed the model skill averaged over all stations. In the following, we examine the skill at individual stations. Figures 4.11 – 4.14 display RMSE maps for all stations in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013. Six plots display snap shots of RMSE maps at forecast hours 24, 48, 72, 96, 120, and 144. Compared with the monthly mean RMSE maps (Figure 4.8), these maps illustrate how RMSE changes with forecast hour.

The RMSE maps for individual forecast hours reveal results consistent with the monthly means (Figure 4.8). The level of model skill follows a trend of gradual reduction with the increase of forecast hour. However, the range of RMSE appears to be confined within a limited range of variation. Across the entire forecast cycle (from hours 24 to 144), RMSE ranged between 1 and 2.5 °C in October, between 1 and 2.0 °C in January, between 1.5 and 3.0 °C in April, and between 2.0 and 5.0 °C in July.

Model skill at individual stations varied over the four months. Model skill was highest in October and January. Model skill was worst in July, and somewhere in between for April. This assessment applies to all stations. Similar to the monthly mean results, the RMSE map at each forecast hour displays a particular modal pattern. Take the results at hour 48 (see Figures 4.11(b) - 14(b)) as an example. The RMSE in October ranged between 1 and 2.5 °C and in January between 1 and 2.0 °C without evident spatial variations. It displays a two-mode pattern in April (Figure 4.12(b)). The RMSE at northern stations (located to the north of 42 °N) were all close to 1 °C, while those to the south were clustered around 2 – 3 °C. The RMSE in July exhibits a

three-mode pattern (Figure 4.14(b)); the mid-latitude (between 36 °N and 45 °N) stations show a RMSE $\sim 3.5\text{--}4.5$ °C, while those to both the south and north exhibit a much smaller RMSE $\sim 1.5\text{--}2.5$ °C.

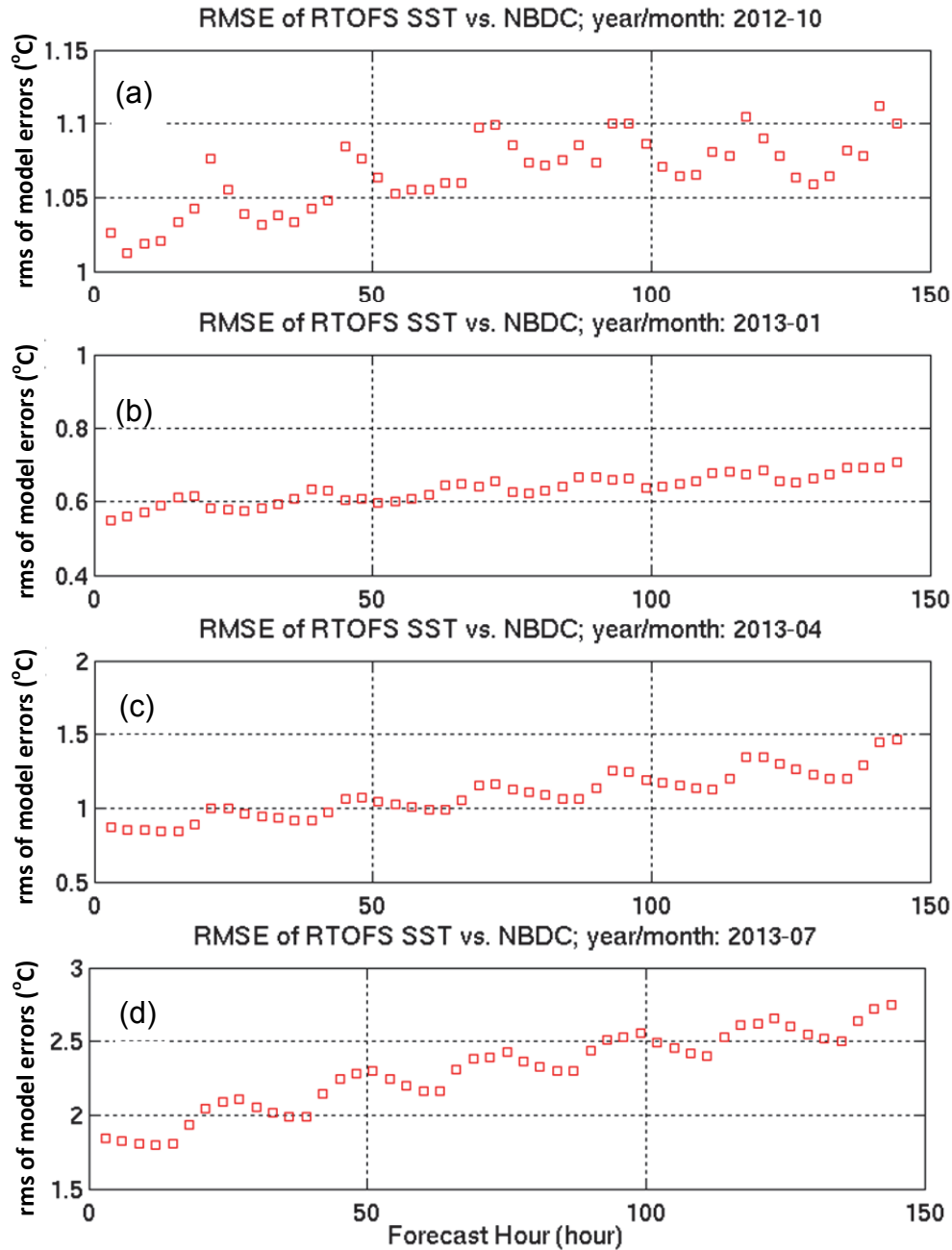


Figure 4.10. Root mean squared errors (RMSE) of the G-RTOFS SST forecast at forecast hours 1-144 in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013.

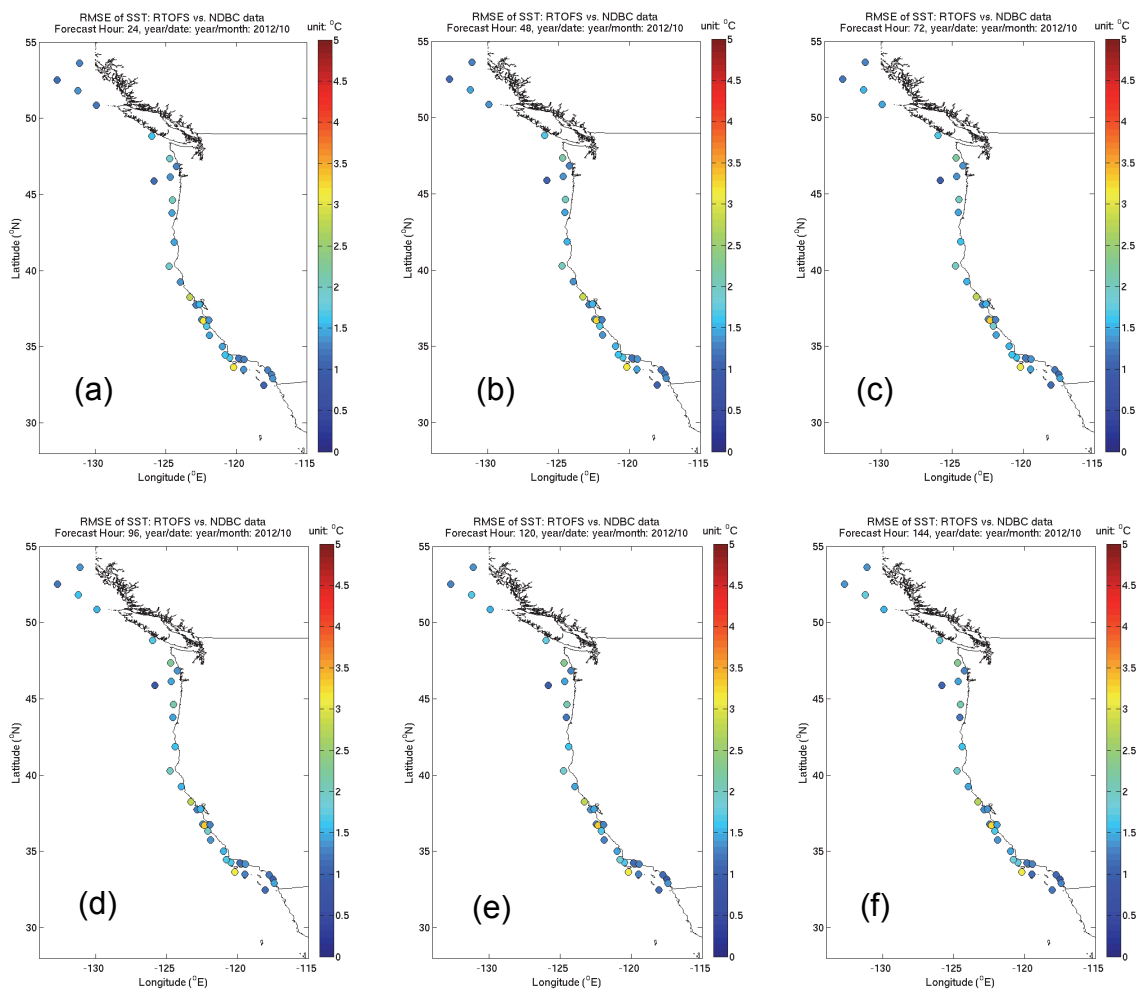


Figure 4.11. Color coded RMSE maps of the G-RTOFS SST forecast in October 2012. The model-data comparison was made against SST measured at 34 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144.

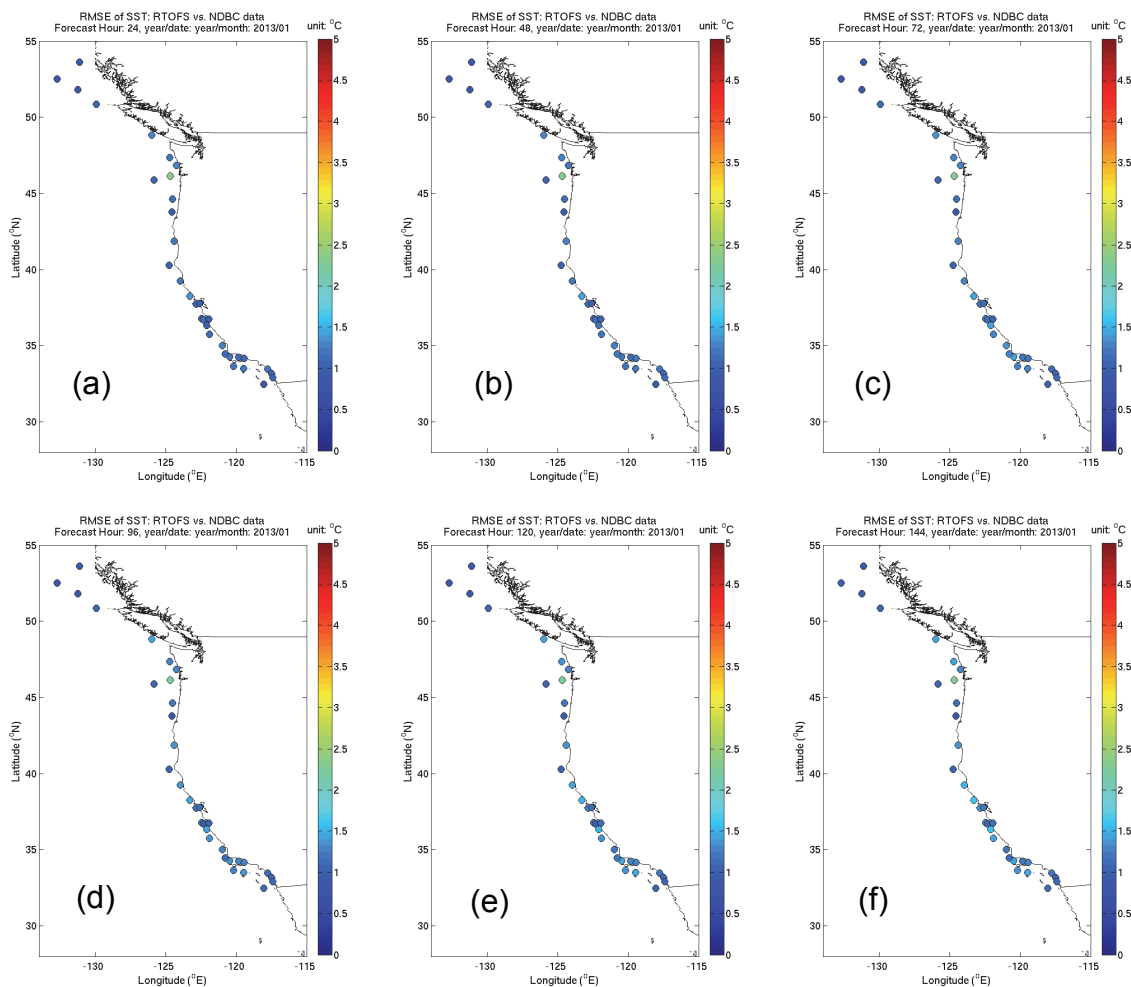


Figure 4.12. Color coded RMSE maps of the G-RTOFS SST forecast in January 2013. The model-data comparison was made against SST measured at 34 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144.

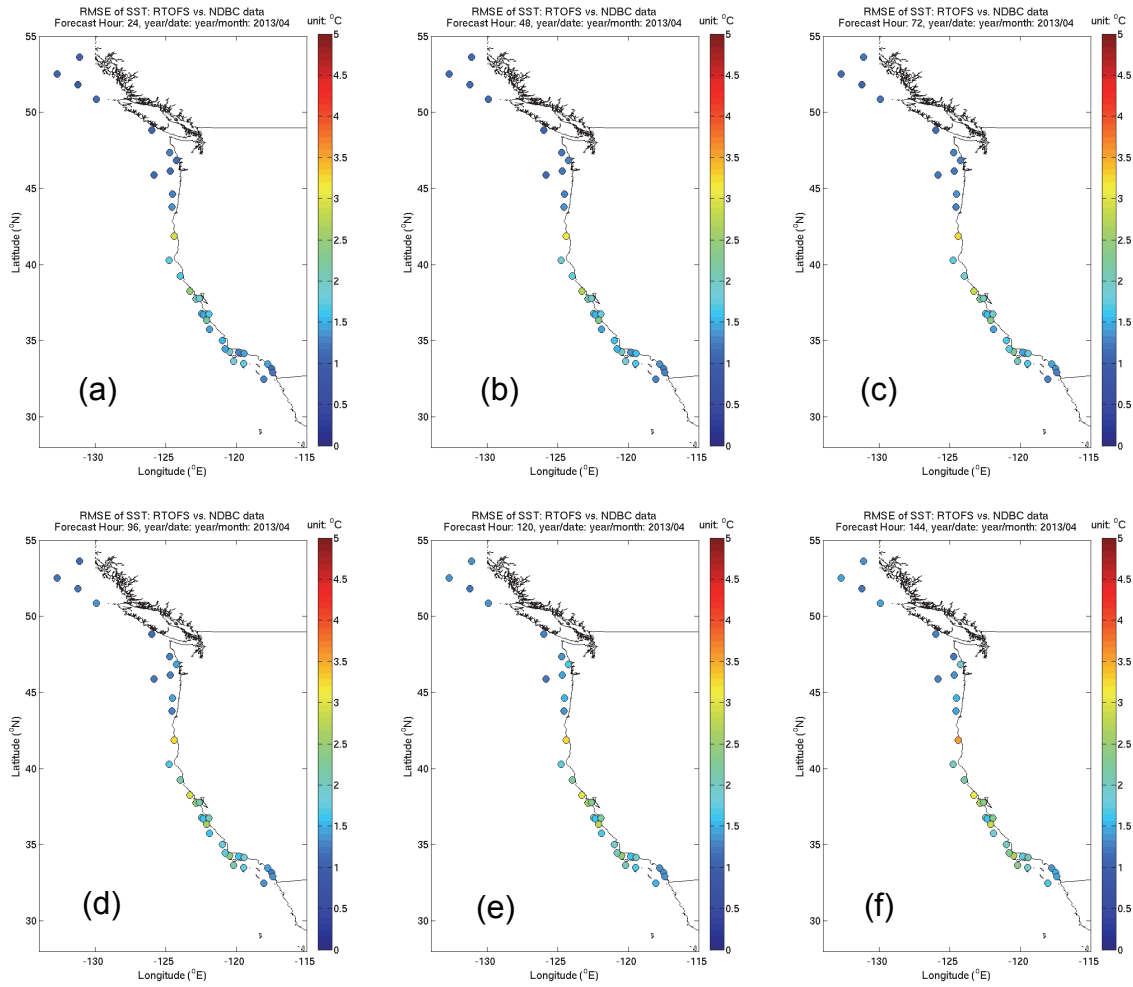


Figure 4.13. Color coded RMSE maps of the G-RTOFS SST forecast in April 2013. The model-data comparison was made against SST measured at 34 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144.

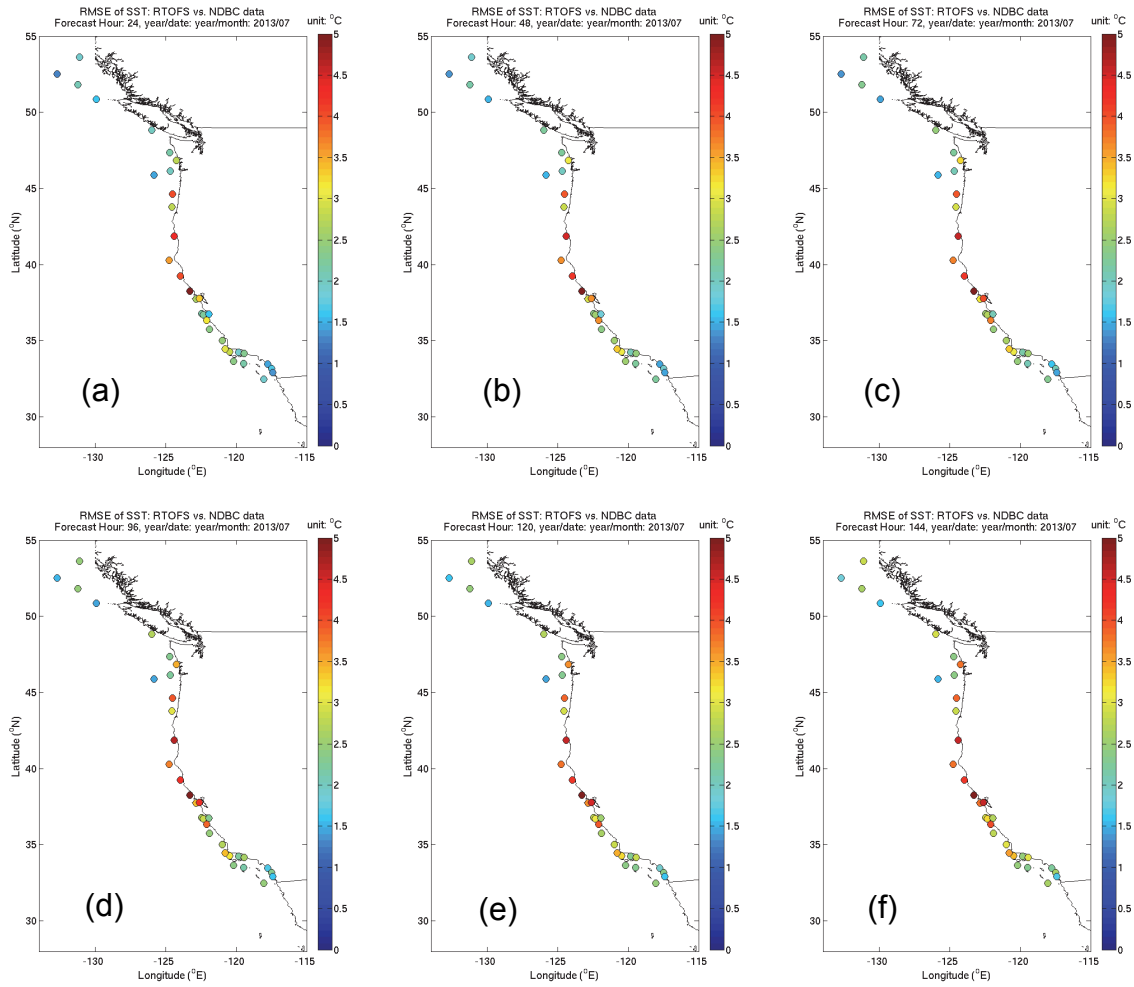


Figure 4.14. Color coded RMSE maps of the G-RTOFS SST forecast in July 2013. The model-data comparison was made against SST measured at 34 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 24, (b) 48, (c) 72, (d) 96, (e) 120, and (f) 144.

4.3. Compared with WOA09 Data

In the following, we compare the G-RTOFS and the WOA09 SST using the same methods, the forecast cycle (FC) based method and forecast hour (FH) based method. Since the WOA09 represents climatological monthly means of oceanic quantities, any comparisons with this database reflect model skills in a climatological sense. Following this line of thinking, we focus on assessing the monthly mean model skills.

4.3.1. FC Based Assessment

Figures 4.15(a) - (d) display maps of the model root-mean-square error (RMSE) for October 2012, January, April, and July 2013. The model exhibits much greater RMSE in July 2013 than in the non-summer months. RMSE was typically less than 2 °C in October, less than 1.5 °C in January, less than 2 °C in April, and around 3-6.5 °C in July.

RMSE was rather evenly valued, 0.5 - 2 °C, across all stations in October, January and April. It displayed large spatial variability in July 2013; stations to the north of 43 °C had RMSE ~ 6.5 °C which was much larger than the RMSE ~ 3-4 °C at the southern stations. This indicates that in a climatological sense the G-RTOFS had much better skill in fall, winter, and spring than in summer.

Figures 4.16(a) - (d) display the mean and standard deviation of model errors at 28 WOA09 data points. Table E.1 lists the geographical locations in longitude and latitude for the 28 WOA09 data points. The mean model-data difference ranges between -1 and 2 °C in October, January, and April without evident bias to either positive or negative values. However, in July, it ranges from 2 °C to 6.5 °C with a visible positive bias. This indicates that the G-RTOFS had better skill in fall, winter, and spring than in summer where it over-predicted the SST.

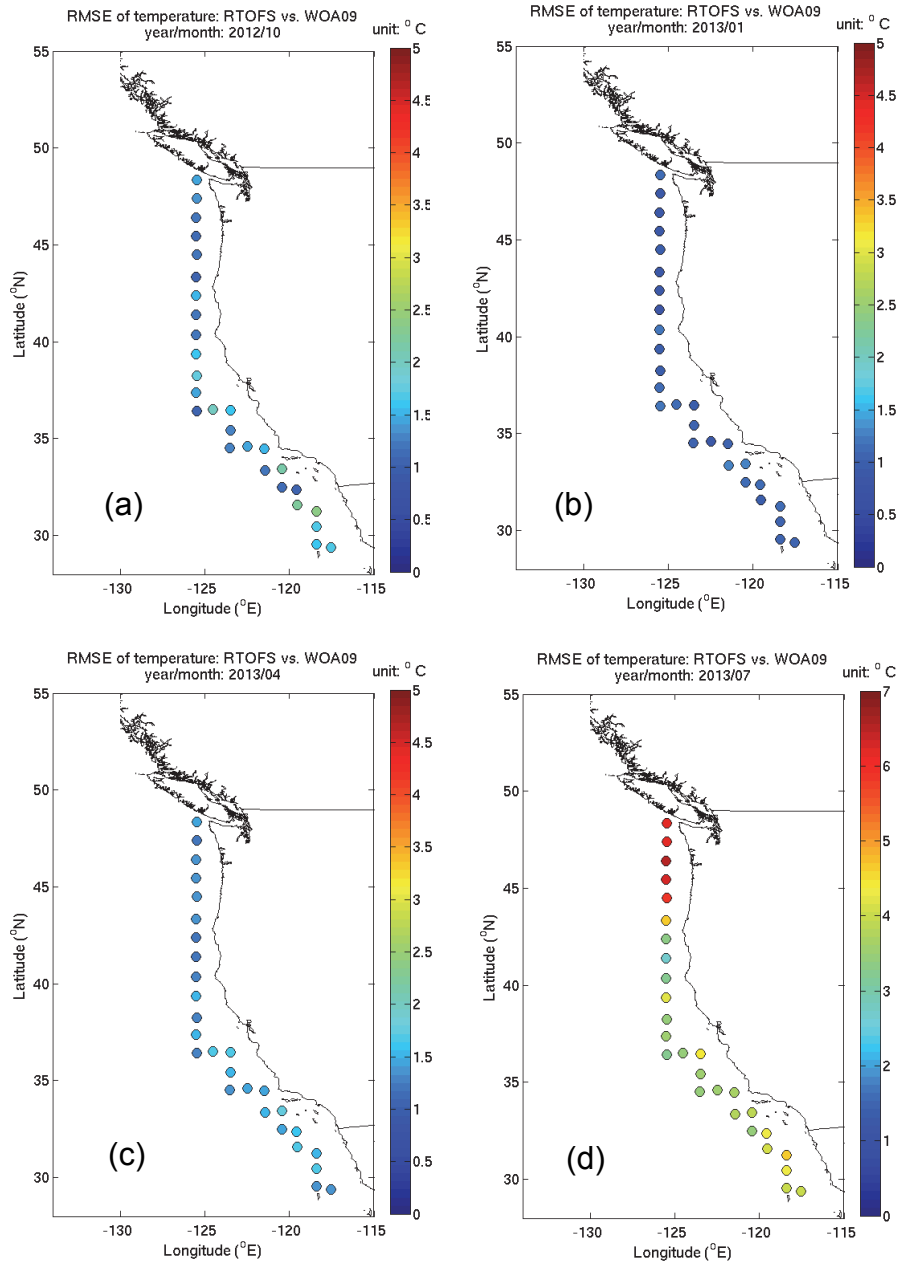


Figure 4.15. Color coded RMSE maps of the G-RTOFS SST forecast in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 28 WOA09 data grid points (Appendix E).

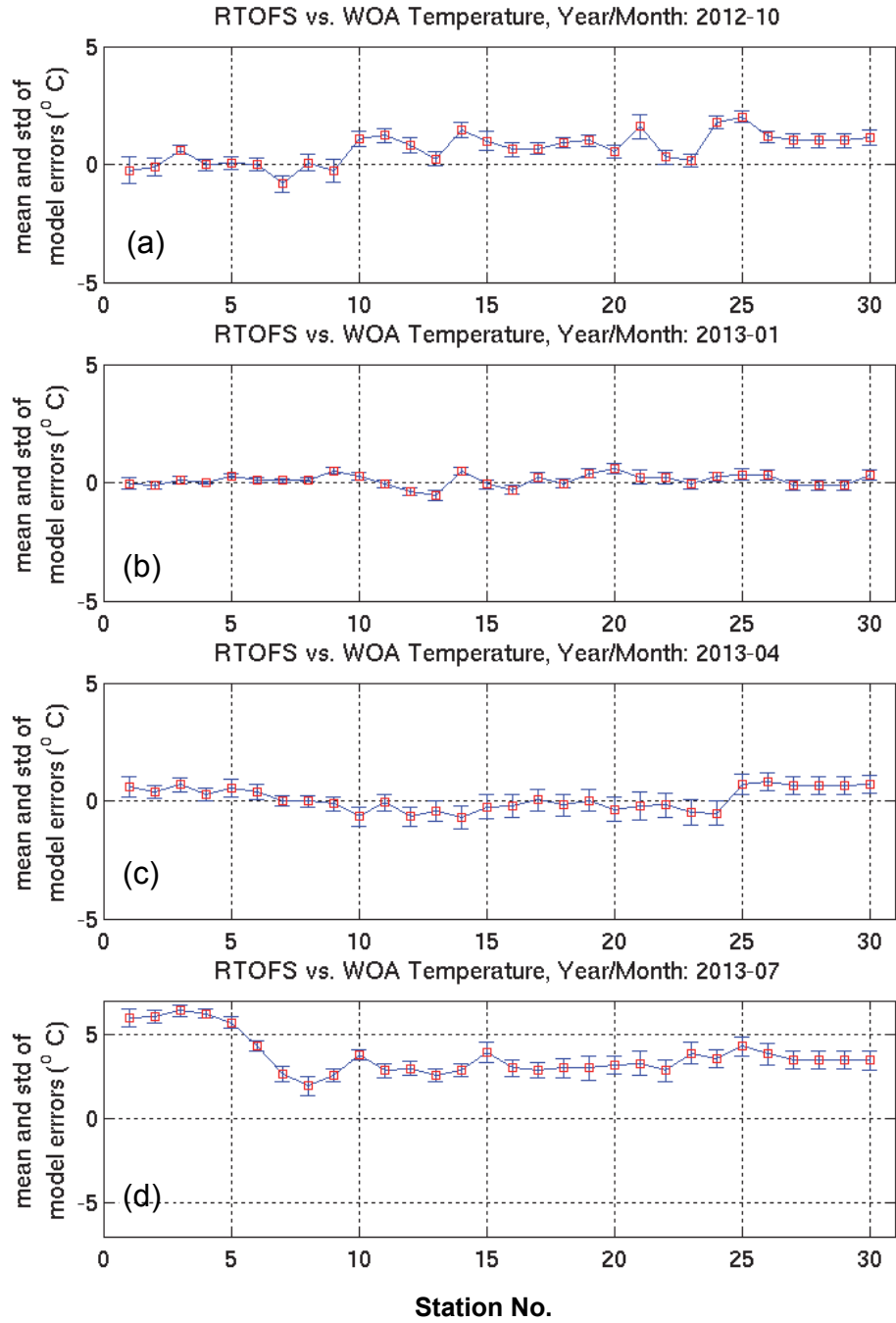


Figure 4.16. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SST errors in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 28 WOA09 data grid points (Table E.1).

4.3.2. FH Based Assessment

In this section, we discuss the evolution of model performance with increasing forecast hour. Here we focus on assessing the model skill in terms of the station-averaged RMSE.

Figures 4.17(a) - (d) display RMSE averaged over 28 selected WOA09 data points at all forecast hours in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013. RMSE was ~ 1 $^{\circ}\text{C}$ across the entire forecast cycle in October, slightly smaller ~ 0.5 $^{\circ}\text{C}$ in January, between 1 $^{\circ}\text{C}$ and 1.5 $^{\circ}\text{C}$ in April, and between 3.8 - 5 $^{\circ}\text{C}$ in July.

In each month, RMSE remained largely invariable throughout the forecast cycle. This is in contrast to the case revealed in the real-time model-data comparison (Sections 4.1.2 and 4.2.2), where the model skill gradually drops with increasing forecast hour. In the climatological sense, the model skill remains steady throughout the forecast cycle.

Note that the RMSE exhibits a regular, diurnal period, sinusoidal perturbation. Of the four months, the signal appears weakest (less than 0.1 $^{\circ}\text{C}$) in January, about 0.2 $^{\circ}\text{C}$ in October and April, and nearly 0.5 $^{\circ}\text{C}$ in July. The diurnal signal evidently resulted from a numerical alias due to the low sampling rate (which was daily in this case) of the time series compiled for the present data analysis, rather than due to any realistic model-data discrepancy.

4.4. Summary

We evaluated the performance of the G-RTOFS SST forecast guidance by comparing the model results with observed data from CO-OPS weather stations, NDBC buoys, and the WOA09 database. The RMSE from all three comparisons reveals similar magnitudes and seasonal variability of model error.

The model skill is more satisfactory in fall, spring, and winter than in summer. Table 4.1 lists the station averaged bias, standard deviation, and root-mean-squared error for the four seasons. In general, all three parameters display seasonal variability. Their magnitudes remain of similar order through fall, winter, and spring, and appear to be noticeably greater in summer. For instance, the bias is around 2-3 $^{\circ}\text{C}$ in summer and less than 1 $^{\circ}\text{C}$ in the other seasons. Similarly, RMSE ranges from 3.3 to 3.8 $^{\circ}\text{C}$ in summer and is less than 1.3 $^{\circ}\text{C}$ in the other seasons. STD exhibits smaller seasonal variability. However, its summer value is still about 0.2-0.9 $^{\circ}\text{C}$ greater than its average value of the other three seasons. Table 4.2 lists the station averaged RMSE of the four seasons estimated using the forecast-hour based method.

Model skill does not appear to degrade as the forecast hour projects into the future. This result differs from the case of the G-RTOFS SSH forecast, where model skill gradually degrades with increasing forecast hour (Section 4.3).

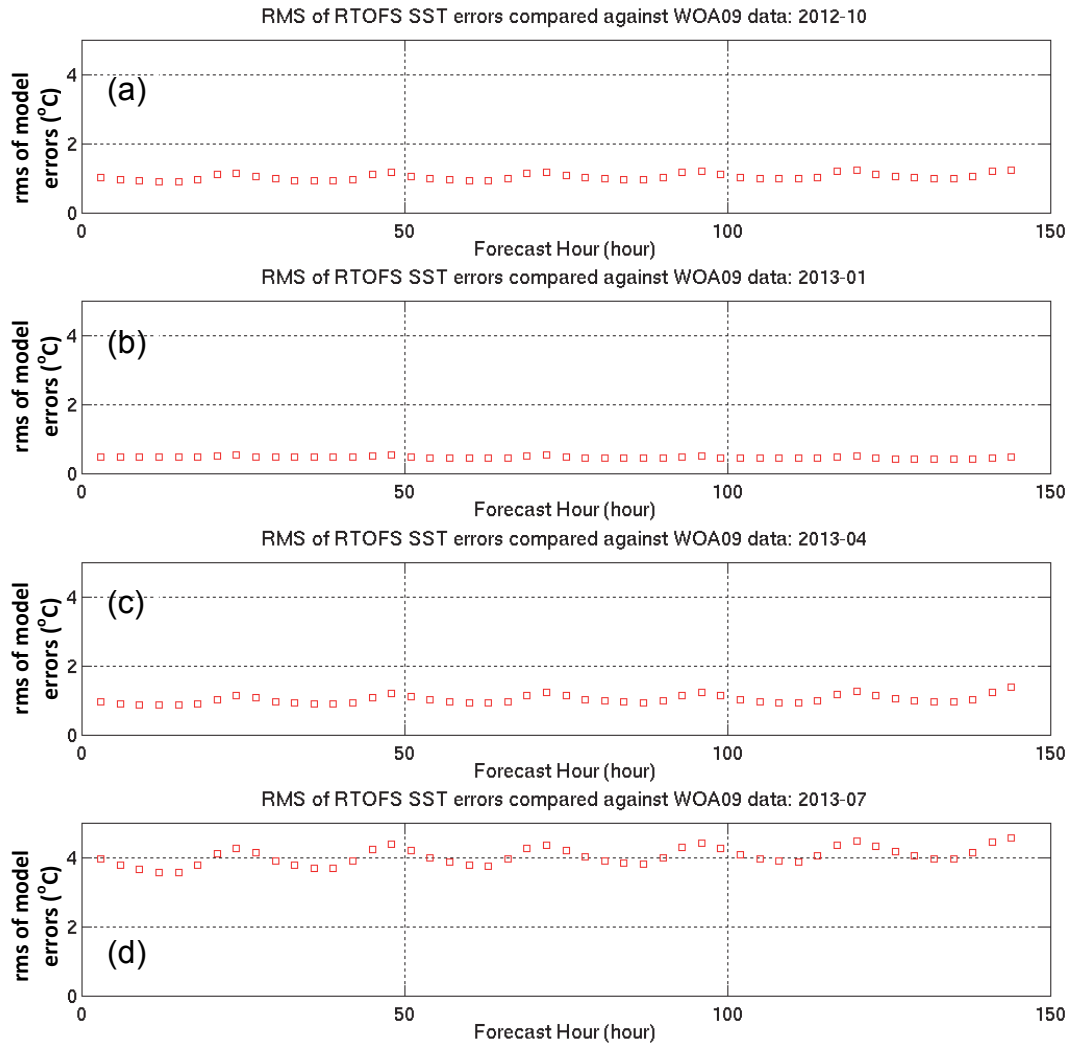


Figure 4.17. Root mean squared errors (RMSE) of the G-RTOFS SST forecast at forecast hours 1-144 in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 28 WOA09 data grid points (Table E.1).

Table 4.1 Forecast cycle-based statistics of the G-RTOFS SST errors.

Data sources	CO-OPS			NDBC			WOA09		
	bias (°C)	std (°C)	rmse (°C)	bias (°C)	std (°C)	rmse (°C)	bias (°C)	std (°C)	rmse (°C)
Fall (Oct. 2012)	0.6	0.6	1.0	0.6	0.5	1.0	0.7	0.3	1.0
Winter (Jan. 2013)	0.6	0.4	0.9	0.2	0.3	0.6	0.1	0.2	0.4
Spring (April 2013)	0.8	0.8	1.3	0.6	0.5	1.0	0.1	0.4	0.9
Summer (July 2013)	2.7	1.5	3.3	1.9	0.7	2.2	3.7	0.5	3.8

Table 4.2 Forecast hour-based RMSE of the G-RTOFS SST.

	RMSE _{CO-OPS} (°C)	RMSE _{NDBC} (°C)	RMSE _{WOA09} (°C)
Fall (Oct. 2012)	1.0-1.2	1.0-1.1	0.9-1.1
Winter (Jan. 2013)	1.0-1.2	0.5-0.7	0.5-0.6
Spring (April 2013)	1.1-1.9	0.8-1.5	1.0-1.1
Summer (July 2013)	2.0-4.1	1.8-2.8	3.8-4.2

Note: RMSE_{CO-OPS}, RMSE_{NDBC}, and RMSE_{WOA09} denote root-mean-squared errors estimated from comparing G-RTOFS SST with observations of CO-OPS weather stations, NDBC buoys, and in WOA09 database, respectively.

5. PERFORMANCE OF G-RTOFS FOR SEA-SURFACE SALINITY

We will evaluate the performance of the G-RTOFS forecast with respect to the sea-surface salinity (SSS). Due to the lack of real time in-situ measurements in the area of present interest, we will focus on comparing the G-RTOFS forecast with the SSS in the WOA09 database. Table E.1 in Appendix E lists the geographical locations, in longitude and latitude, of the 28 WOA09 stations.

We used the two same methods as those described in previous chapters, namely, the forecast cycle (FC) based and forecast hour (FH) based methods. Here, we used the climatological monthly mean SSS in WOA09. Hence, any comparisons with this database reflect the model skills in a climatological sense. In the following, we focus on assessing the monthly mean model skills. Since we are comparing the model results with climatological data sets, we do not think it would make scientific sense to address at length details about the forecast hour based model performance. In the following, we investigate only the station averaged forecast cycle based model skills so as to attain a general insight into the model skill.

5.1. Forecast Cycle Based Assessment

Figures 5.1(a) - (d) show the maps of the model root-mean-square error (RMSE) in October 2012, January, April, and July 2013. RMSE was typically less than 0.7 psu in all four months. However, the RMSE varied in both spatially and in magnitude across the months. It ranged from 0.1 psu to 0.5 psu over the entire domain in October. In January and April, the RMSE was slightly elevated at two to three stations near the Strait of Juan de Fuca. In July 2013, the RMSE ranged between 0.1 psu and 0.7 psu with greater spatial variation than the other months. In sharp contrast to the spatially homogeneous RMSE map in October, the RMSE in July demonstrated evident spatial variability with sporadically greater RMSE at stations near the entrance to the Strait of Juan de Fuca, between 42 °N and 45 °N, and at stations between 34 °N and 35.5 °N. In a climatological sense, the G-RTOFS demonstrated better skill in fall, winter, and spring than in summer.

Figures 5.2(a) - (d) display the mean and standard deviation of model errors at 28 WOA09 stations. The mean model-data difference typically ranges between -0.3 psu and 0.4 psu in October, January, and April with a slight positive bias. The mean model-data difference ranges between -0.5 psu and -0.3 psu in July.

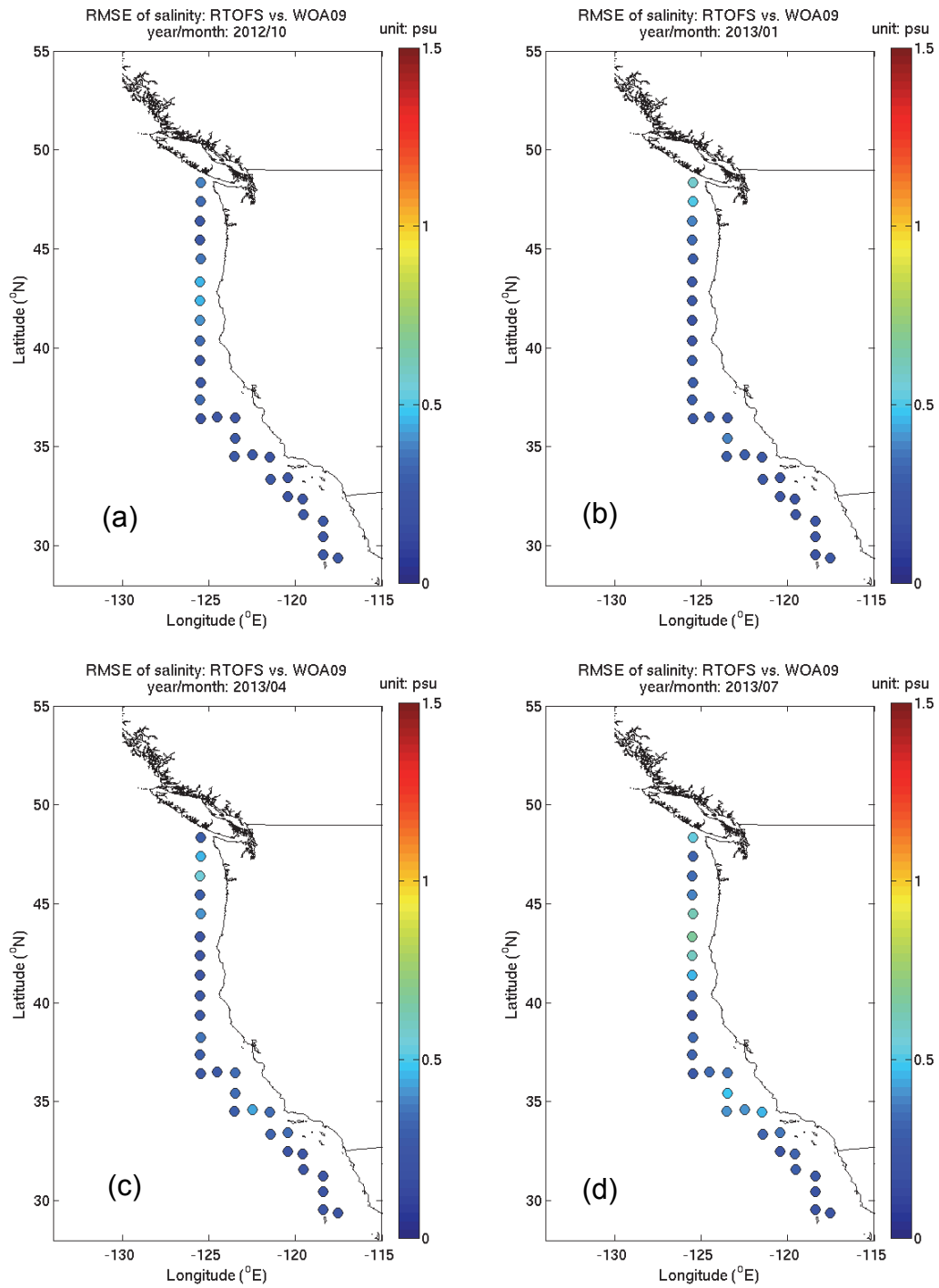


Figure 5.1. Color coded RMSE maps of the G-RTOFS SSS forecast in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 28 WOA09 data grid points (Table E.1).

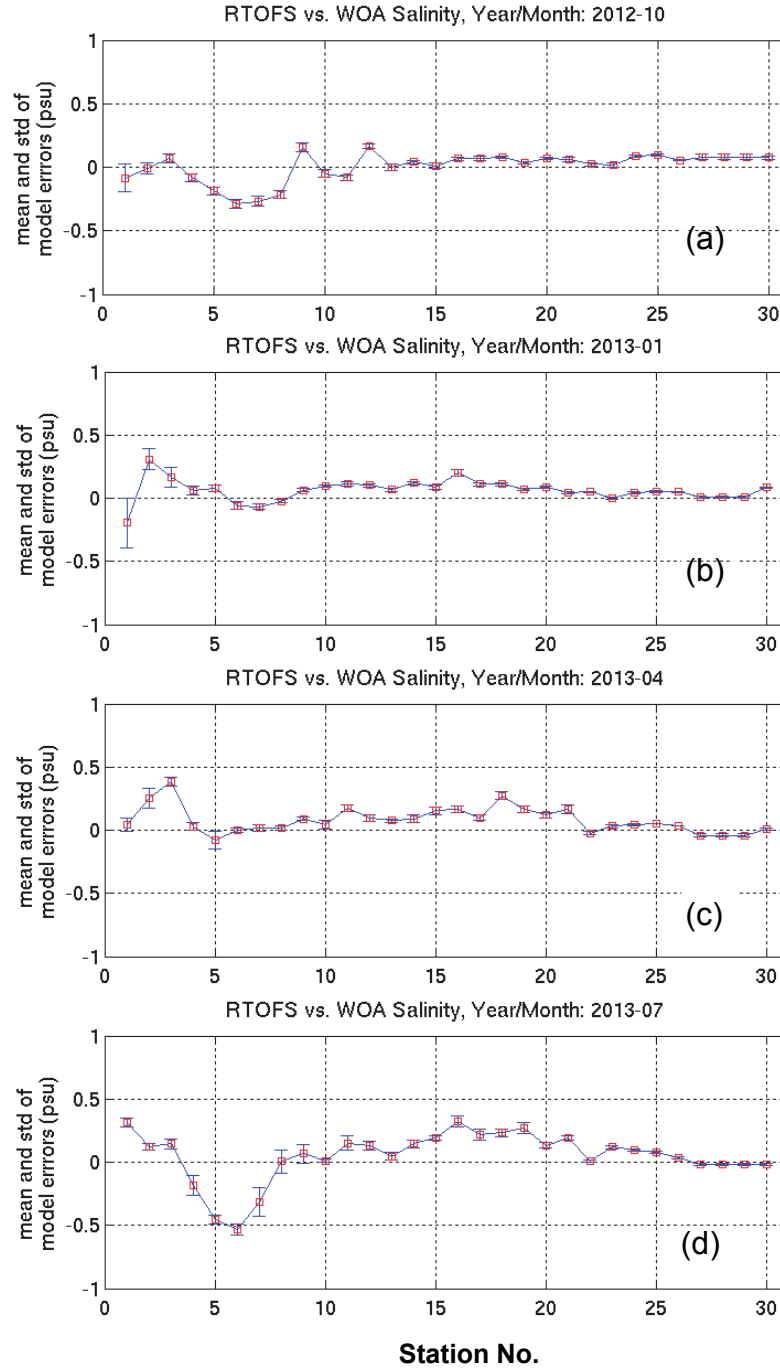


Figure 5.2. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SSS errors in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 28 WOA09 data grid points (Table E.1).

5.2. Forecast Hour Based Assessment

In this section, we discuss the dependence of model performance with respect to increasing forecast hour. For the reason stated in the opening paragraph of Section 4.3, we will focus on assessing the model skill in terms of the station-averaged RMSE.

Figures 5.3(a) - (d) display RMSE averaged over 28 WOA09 stations at all 144 forecast hours in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013. Remarkably, RMSE remained at a nearly constant level throughout the entire forecast cycle in each month. The RMSE was 0.2 psu in October, 0.1 psu in January and April, and 0.2 psu in July.

In contrast to the cases for SSH (Section 3.3) and SST (Section 4.3), the plots did not reveal any evident diurnal period signals, even though the SSS analysis adopted the same data sampling procedures as those for the SSH and SST. We attribute this to the inertial nature of SSS. The SSS from the RTOFS forecast were nearly constant throughout the forecast cycle. We therefore would not expect to observe any significant diurnal variations from the RMSE plots.

5.3. Summary

The G-RTOFS SSS forecast guidance demonstrates a satisfactory agreement with the WOA09 database. The station averaged bias ranges between -0.3 psu and 0.4 psu in the fall, winter, and spring and spreads over a slightly broader range, from -0.5 psu to -0.3 psu in the summer. The STD is less than 0.2 psu throughout all four seasons. The RMSE remains at a nearly constant level of about 0.2 psu throughout a forecast cycle across all four seasons.

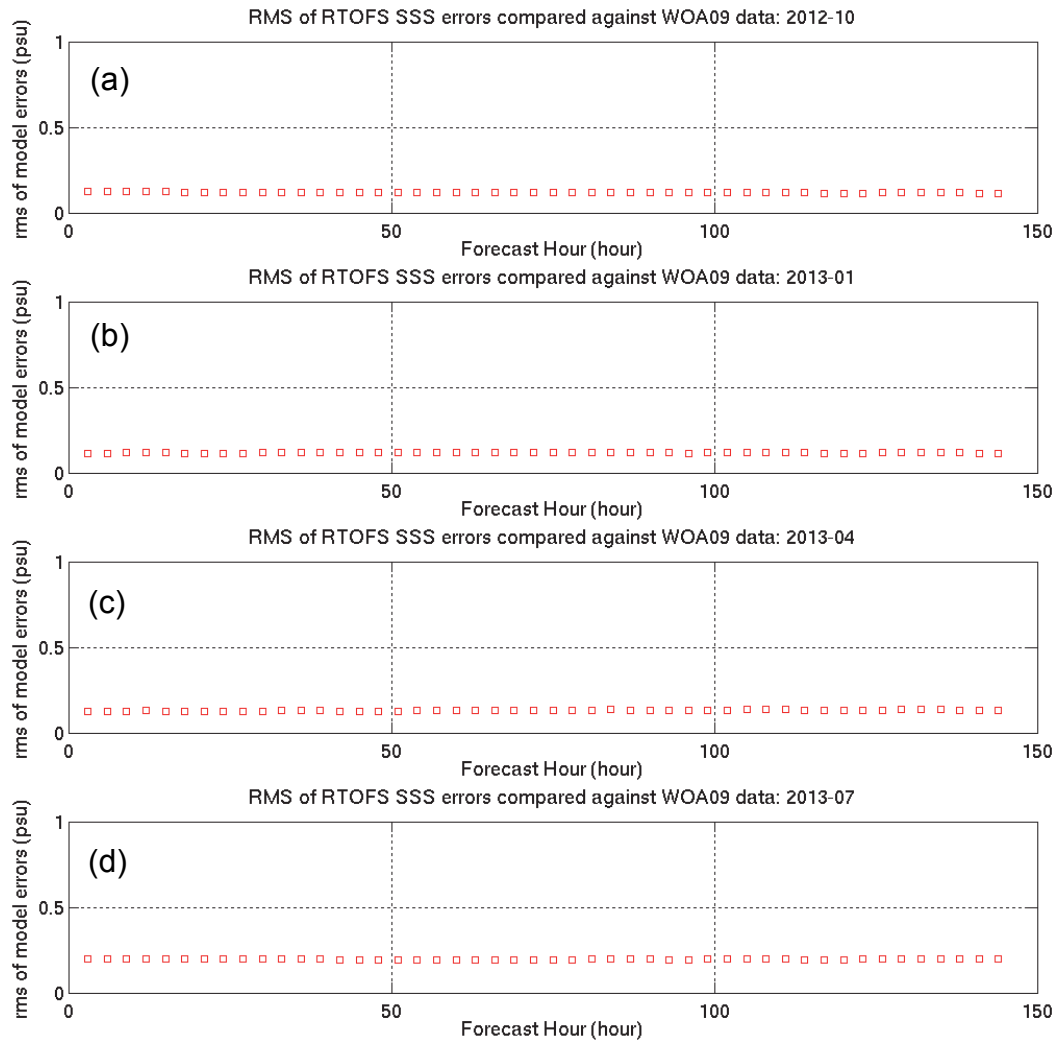


Figure 5.3. Root mean squared error (RMSE) of the G-RTOFS SSS forecast at forecast hours 1-144 in (a) October 2012, (b) January 2013, (c) April 2013, and (d) July 2013 at 28 WOA09 data grid points (Table E.1).

6. SUMMARY

The Operational Nowcast/Forecast Systems (OFS) of NOAA/NOS use sea-surface height, temperature, and salinity outputs from the G-RTOFS in preparing open ocean boundary forcing conditions, as well as the sea-surface height from the ETSS when G-RTOFS SSH is not available. To support future development of NOS OFS in western U.S. coastal waters, we assessed the performance of SSH forecasts from both the G-RTOFS and ETSS, and the forecasts of sea-surface temperature (SST) and sea-surface salinity (SSS) from the G-RTOFS.

This report summarizes the procedures and results of assessment. Section 1 describes background information concerning the set up and operation of both the G-RTOFS and ETSS, as well as the observed data (water level, SST, and SSS) used in model-data comparisons. Chapter 2 explains technical details of the two methods used to assess model performance: the forecast cycle (FC) based method and the forecast hour (FH) based method. Chapters 3-5 discuss the model performance of the SSH (from the G-RTOFS and ETSS), SST (from G-RTOFS), and SSS (from G-RTOFS) forecast guidance, respectively.

A purpose of this study is to investigate the seasonal variability of model performance. We chose four monthly periods: October 2012 and January, April, and July of 2013. These four months represent the seasons of fall, winter, spring, and summer, respectively.

To evaluate the SSH performance, quality controlled hourly time series measured at CO-OPS water level stations (Table A.1 and Table B.1) were used. We first low-pass filtered the time series to remove their tidal components to obtain a sub-tidal water level time series. We then compared the subtidal water level time series with the model results.

We used SST measurements from 13 CO-OPS meteorological observation stations (Table C.1) and 34 NDBC buoys (Table D.1) to evaluate the performance of the G-RTOFS SST forecast. The NDBC stations are more often located in offshore regions than are the CO-OPS stations. Hence, the NDBC data are more representative of the offshore SST fields. In addition, we used SST/SSS values at 30 nearshore grid points (Table E.1) in the WOA09 database to assess the performance of the G-RTOFS SST/SSS in a climatological sense. In future work, it would be worthwhile to investigate the performance of the G-RTOFS subsurface temperature and salinity vs. the WOA09.

We assessed the model performance from two time-frame perspectives. One is across each forecast cycle and the other is with respect to the forecast hour. The two perspectives correspond to the methods described in Chapter 2, namely, the forecast cycle (FC) based method and the forecast hour (FH) based method. The FC method calculates bias and the root-mean-square error (RMSE) across each cycle. Note that the G-RTOFS generates a forecast cycle four times per day at hours 00z, 06z, 12z, and 18z UTC. The ETSS forecast is generated one time per day at hour 00z UTC. This study focuses on analyzing the model results produced on the 00z UTC cycle.

The FH method assesses how the model performance evolves as a forecast progresses in a cycle. This method calculates bias and the RMSE of model results at each forecast hour over multiple

forecast cycles. We analyzed the three-hourly G-RTOFS SSH/SST/SSS forecast guidance from hour 00 to hour 144 UTC and hourly ETSS SSH guidance from hour 00 to hour 96 UTC.

In general, the forecast guidance of SSH (from the G-RTOFS and ETSS) and SST/SSS (from the G-RTOFS) demonstrates satisfactory agreement with observations. The level of model performance varies by seasons and by the forecast hour.

When using the FC method to assess SSH forecast guidance from both models, the STD ranges between 2 and 4 cm for ETSS and between 2 and 10 cm for G-RTOFS through all four seasons. The bias of ETSS is about -7 cm in the fall and winter and ranges from 0.1 to 2.2 cm in the spring and summer. By contrast, the G-RTOFS displays a bias ranging between 0 and 18 cm over all four seasons. The RMSE for ETSS increases from 6 to 11 cm in the spring and summer to ~13 cm in the fall and winter. The RMSE for G-RTOFS remains rather steady year round with a maximum RMSE of 16 cm, 19 cm, 20 cm, and 15 cm for October, January, April, and July, respectively.

With respect to the forecast hour, the skill of both models degrades as a forecast progresses through the forecast cycle. The RMSE of ETSS increases from less than 2 cm at the beginning of a forecast cycle to around 6 to 10 cm near the end of the cycle. The RMSE of the G-RTOFS SSH forecast increases monotonically from less than 2 cm at the start of a forecast cycle to a plateau from the mid-point of the cycle until the end. Model skill exhibits seasonal variability. The maximum RMSE for ETSS is 8 cm, 10 cm, 10 cm, and 6 cm in fall, winter, spring, and summer, respectively. The maximum RMSE_{G-RTOFS} are 6.5 cm, 11.5 cm, 10.0 cm, and 4.75 cm which correspond to the fall, winter, spring, and summer, respectively.

The performance of G-RTOFS SST forecast guidance demonstrates seasonal variability. The performance is more satisfactory in the fall, winter, and spring than in summer. Station averaged bias, STD, and RMSE remain of a similar magnitude in the fall, winter, and spring, and are noticeably greater in the summer. For instance, the bias ranges between 2 and 3 °C in the summer and is less than 1 °C in the other seasons. The RMSE is about 3.3 to 3.8 °C in the summer and less than 1.3 °C in the other seasons. The STD exhibits smaller seasonal variability. However, its summer value is still about 0.2 – 0.9 °C greater than its average magnitude in the other three seasons. The model skill does not degrade with increasing forecast hour of each forecast cycle. RMSE remains nearly invariable throughout each forecast cycle as opposed to the gradual skill degradation which occurs in case of the G-RTOFS SSH forecast.

The SSS forecast guidance from the G-RTOFS demonstrates a satisfactory agreement with the WOA09 database. The station averaged bias ranges between -0.3 psu and 0.4 psu in the fall, winter, and spring. It spreads over a slightly broader range between -0.5 psu and -0.3 psu in the summer. The STD is less than 0.2 psu throughout all four seasons. The RMSE remains at a nearly constant level of 0.2 to 0.4 psu through a forecast cycle across all four seasons.

ACKNOWLEDGEMENTS

We are very grateful to Dr. Frank Aikman, Chief of Marine Modeling and Analysis Programs, Coastal Survey Development Laboratory (CSDL) for providing the project guidance and critical resources. We thank Dr. Avichal Mehra, National Weather Service's Environmental Modeling Center, for providing guidance on how to retrieve G-RTOFS forecast data archived in the NCEP's High Performance Storage System (HPSS). We wouldn't have been able to complete this project without their gracious help and encouragement.

REFERENCES

Global Real-Time Ocean Forecast System (G-RTOFS), 2011. <http://polar.ncep.noaa.gov/global>.

Metzger, E.J., O.M. Smedstad, P.G. Thoppil, H.E. Hurlburt, J.A. Cummings, A.J. Wallcraft, L. Zamudio, D.S. Franklin, P.G. Posey, M.W. Phelps, P.J. Hogan, F.L. Bub, and C.J. DeHaan. 2014. US Navy operational global ocean and Arctic ice prediction systems. *Oceanography* 27(3):32–43.

NOAA/NCEP's Extra Tropic Storm Surge system (ETSS), 1992. <http://nws.weather.gov/mdl/surge/etsurge2.0>.

Appendix A. CO-OPS Water Level Stations Used for the ETSS Skill Assessment

Table A.1. Water Level Stations Used in ETSS Skill Assessment

No.	Station ID	Station Name	Longitude (°E)	Latitude (°N)
1	9414290	San Francisco, CA	-122.465	37.807
2	9415020	Point Reyes, CA	-122.977	37.996
3	9416841	Arena Cove, CA	-123.708	38.913
4	9418767	North Spit. Humboldt Bay, CA	-124.217	40.767
5	9419750	Crescent City, CA	-124.183	41.745
6	9431647	Port Orford, OR	-124.498	42.739
7	9432780	Charleston, OR	-124.322	43.345
8	9435380	South Beach. Yaquina River, OR	-124.043	44.625
9	9439011	Westport, WA	-124.105	46.904
10	9439040	La Push, WA	-124.637	47.913
11	9440910	Neah Bay, WA	-124.612	48.367
12	9443090	Toke Point, WA	-123.967	46.707

Appendix B. CO-OPS Water Level Stations Used for the G-RTOFS Skill Assessment

Table B.1. CO-OPS water level stations Used in G-RTOFS water level skill assessment

No.	Station ID	Station Name	Longitude (°E)	Latitude (°N)
1	9410230	La Jolla, CA	-117.258	32.867
2	9410660	Los Angeles, CA	-118.272	33.72
3	9410840	Santa Monica, CA	-118.5	34.008
4	9411340	Santa Barbara, CA	-119.685	34.408
5	9412110	Port San Luis, CA	-120.76	35.177
6	9413450	Monterey, CA	-121.888	36.605
7	9414290	San Francisco, CA	-122.465	37.807
8	9415020	Point Reyes, CA	-122.977	37.996
9	9416841	Arena Cove, CA	-123.708	38.913
10	9418767	North Spit. Humboldt Bay, CA	-124.217	40.767
11	9419750	Crescent City, CA	-124.183	41.745
12	9431647	Port Orford, OR	-124.498	42.739
13	9432780	Charleston, OR	-124.322	43.345
14	9435380	South Beach, Yaquina River, OR	-124.043	44.625
15	9437540	Garibaldi, OR	-123.956	45.57
16	9439011	Westport, WA	-124.105	46.904
17	9439040	La Push, WA	-124.637	47.913
18	9440910	Neah Bay, WA	-124.612	48.367

Appendix C. CO-OPS Meteorological Observation Stations for Assessing the G-RTOFS SST Performance

Table C.1. CO-OPS stations used in the G-RTOFS SST Skill Assessment

No.	Station ID	Station Name	Longitude (°E)	Latitude (°N)
1	9410230	La Jolla, CA	-117.258	32.867
2	9410840	Santa Monica, CA	-118.5	34.008
3	9412110	Port San Luis	-120.76	35.177
4	9413450	Monterey, CA	-121.888	36.605
5	9415020	Point Reyes, CA	-122.977	37.996
6	9416841	Arena Cove, CA	-123.708	38.913
7	9418767	North Spit. Humboldt Bay, CA	-124.217	40.767
8	9419750	Crescent City, CA	-124.183	41.745
9	9431647	Port Orford, OR	-124.498	42.739
10	9435380	South Beach. Yaquina River, OR	-124.043	44.625
11	9437540	Garibaldi, OR	-123.956	45.57
12	9442396	La Push, WA	-124.637	47.913
13	9443090	Neah Bay, WA	-124.612	48.367

Appendix D. NDBC Buoy Stations for Assessing the G-RTOFS SST Performance

Table D.1. Identification numbers, names, and longitudes, and latitudes of NDBC Buoy Stations

No.	Station ID	Station Name	Longitude (°E)	Latitude (°N)
1	46011	Santa Maria - 21 NM Northwest of Point Arguello, CA	-120.992	35
2	46013	Bodega Bay - 48NM North Northwest of San Francisco, CA	-123.301	38.242
3	46014	Pt Arena - 19NM North of Point Arena, CA	-123.974	39.235
4	46026	San Francisco - 18NM West of San Francisco, CA	-122.833	37.759
5	46027	St Georges - 8NM West Northwest of Crescent City, CA	-124.381	41.85
6	46028	Cape San Martin - 55NM West Northwest of Morro Bay, CA	-121.884	35.741
7	46041	Cape Elizabeth- 45NM Northwest of Aberdeen, WA	-124.731	47.353
8	46042	Monterey - 27NM West of Monterey Bay, CA	-122.469	36.785
9	46050	Stonewall Banks - 20NM West of Newport, OR	-124.534	44.639
10	46053	E. Santa Barbara - 12NM Southwest of Santa Barbara, CA	-119.841	34.248
11	46054	Santa Barbara W 38 NM West of Santa Barbara, CA	-120.459	34.274
12	46069	South Santa Rosa Island, CA	-120.2	33.67
13	46086	San Clemente Basin, Ca	-118.034	32.491
14	46089	Tillamook, OR - 85 NM WNW of Tillamook, OR	-125.83	45.889
15	46092	MBM1, CA	-122.02	36.75
16	46114	West Monterey Bay, CA	-122.342	36.71
17	46147	South Moresby, WA	-131.22	51.83
18	46183	North Hecate Strait, WA	-131.1	53.62
19	46206	La Perouse Bank, WA	-126	48.84
20	46207	East Dellwood, WA	-129.92	50.87
21	46208	West Moresby, WA	-132.69	52.52
22	46211	Grays Harbor, WA (036)	-124.244	46.857
23	46213	Cape Mendocino, CA (094)	-124.74	40.294
24	46217	Anacapa Passage, CA (111)	-119.435	34.167
25	46218	Harvest, CA (071)	-120.769	34.451
26	46223	Dana Point, CA (096)	-117.767	33.458
27	46224	Oceanside Offshore, CA (045)	-117.471	33.179
28	46225	Torrey Pines Outer, CA (100)	-117.393	32.93
29	46229	Umpqua Offshore, OR (139)	-124.551	43.769
30	46236	Monterey Canyon Outer, CA (156)	-121.947	36.761

31	46237	San Francisco Bar, CA (142)	-122.599	37.781
32	46238	San Nicolas Island North, CA (167)	-119.49	33.499
33	46239	Point Sur, CA (157)	-122.101	36.338
34	46248	Astoria Canyon, OR (179)	-124.667	46.133

Appendix E. WOA09 Grid Nodes for Assessing the G-RTOFS SST/SSS Performance

Table E.1. List of longitudes, and latitudes of WOA09 Grid Nodes

No.	Station Name	Longitude (°E)	Latitude (°N)
1	WOA09-Sta 1	-125.4847	48.3396
2	WOA09-Sta 2	-125.4847	47.4124
3	WOA09-Sta 3	-125.5166	46.4079
4	WOA09-Sta 4	-125.5166	45.4421
5	WOA09-Sta 5	-125.4847	44.5149
6	WOA09-Sta 6	-125.5006	43.3558
7	WOA09-Sta 7	-125.5006	42.39
8	WOA09-Sta 8	-125.5006	41.3855
9	WOA09-Sta 9	-125.5166	40.3424
10	WOA09-Sta 10	-125.5006	39.3766
11	WOA09-Sta 11	-125.4847	38.2562
12	WOA09-Sta 12	-125.5006	37.3676
13	WOA09-Sta 13	-125.4847	36.4404
14	WOA09-Sta 14	-124.5123	36.5177
15	WOA09-Sta 15	-123.4762	36.4791
16	WOA09-Sta 16	-123.4762	35.4359
17	WOA09-Sta 17	-123.508	34.5087
18	WOA09-Sta 18	-122.4878	34.586
19	WOA09-Sta 19	-121.4995	34.4701
20	WOA09-Sta 20	-121.4488	33.3588
21	WOA09-Sta 21	-120.4175	33.4471
22	WOA09-Sta 22	-120.4175	32.4758
23	WOA09-Sta 23	-119.5795	32.3875
24	WOA09-Sta 24	-119.515	31.5928
25	WOA09-Sta 25	-118.4192	31.2396
26	WOA09-Sta 26	-118.4192	30.4449
27	WOA09-Sta 27	-118.4192	29.5619
28	WOA09-Sta 30	-117.5503	29.4052